

OM of: US-08-900-220C-17\_COPY\_20\_198 to: Issued\_Patents\_NA.\*

out\_format : pfs

Date: Apr 22, 2001 12:05 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

#### Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xip  
-O=/cgn2\_1/USPTO\_spool/US08900220/runat.19042001.104047.21686/app.query.fasta.1.253  
-DB=Issued\_Patents\_NA -OEMT=fastaf -SUFFIX=rni -GAPOP=12.000  
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.500  
-OGAPOP=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500  
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500  
-DELCP=6.000 -DELCP=7.000 -START=1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFM=pfs  
-NORM=ext -MINLEN=0 -MAXLEN=200000000  
-USER=US08900220 -CGN1\_1\_55 -NCPU=6 -ICPU=3 -LONGLOG -NO\_XLPHY  
-WAIT -THREADS=1

#### Search information block:

Query: US-08-900-220C-17\_COPY\_20\_198  
Query length: 179  
Database: Issued\_Patents\_NA.\*  
Database sequences: 302621  
Database length: 87301344  
Search time (sec): 80.530000

#### score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	..	8e-103	1190
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-9				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-9				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-9				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-9				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-7				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-946-329A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-176-427B-2				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-4				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-356-060A-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-6				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-1				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5A_COMB.seq:US-08-748-591-5				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-460-900C-3				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US-08-176-427B-9				932.00	1990.29	2.8e-103	1190	1
/cgn2_6/prodata/2/ina/5B_COMB.seq:US								

```
|||||
158 CCGAGGCGACCTGGGGCGAGTGGCCACGAGGAGGGAGGTTACAGG 207
51 GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
208 GGGTGCGAGGCGCTCCGGGACCTCGTACCCCACTACACCCGACATAT 257
67 ePhelysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
258 CTTCAGGATGAGGAGACAGACCGCGCAGACCCCTGATGACAGCGTT 307
84 yslYsgIuArgValAsnAlaLeuAlaIleAlaValMetAsnMetPro 100
308 GCNAAGGCGGGTGACGCTGTAGCCATCGCGGATGATGATGTGGCC 357
101 GlyValArgLeuArgValThrGluGlyTrrAspGluAspGlyHisAl 117
358 GGAGTACGCTAGTGTGATGAAAGGTGGGAGAGAGCGCCACACGCG 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThSera 134
408 ACAGATTCACTCCACTACGAGAGCGCGCTTGACATCACACGCTGTG 457
134 sPaRgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
458 ACCGTGACCGCTAATAGTATGTTGTGGCGCGCTAGCTGTGGAAGCC 507
151 GlyPheAspTrrValTyrTyrGluSerArgAsnHisValHisValSera 167
508 GGATTGACAGTGGGTCTACTACGAGTCCCGCACACACATCCACGATG 557
167 llySAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
558 CAAAGCTGATTACTCACTGCGGTCCGAGCGCGAGGC 594

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:us-08-356-060A-2

seq_documentation_block:
; Sequence 2, Application US/08356060A
; Patent No. 5844079
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 14-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMI-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
```

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1190 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1191
US-08-356-060A-2

alignment_scores:
Quality: 932.00 Length: 179
Ratio: 5.207 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.207

alignment_block:
US-08-900-220c-17_copy_20_198 x US-08-356-060A-2 ..
Align seq 1/1 to: US-08-356-060A-2 from: 1 to: 1190

1 AAGlnSerGlySgIyProGlyArgGlyProValGlyArgArgArgTrrAl 17
58 GCCAGAGCTGCGGGCGCGGCGGAGGACCGGTGGCGCGGGGTATGT 107
17 aArgGlyGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
108 GCCGACCAACTGTGCTGCTGCTGATACAGCAGATTGTGCCAGTATGC 157
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
158 CCGAGCGGACCTGCGGCGCGAGTGGCCAGCGGAGGAGGATGAACAG 207
51 GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
208 GGGTGCGAGCGCTTCCGGGACCTGTGATCCCACTACAAACCCGACATA 257
67 ePhelysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
258 CTTCAGGATGAGGAGAGACAGCGCGCGACCGCTGATGACAGAGGTT 307
84 yslYsgIuArgValAsnAlaLeuAlaIleAlaValMetAsnMetPro 100
308 GCAAGAGCGCGTGAACCGCTCTAGCCATCCCGGTGATGACATGTGCG 357
101 GlyValArgLeuArgValThrGluGlyTrrAspGluAspGlyHisAl 117
358 GGAGTACGCTTACGTGTGACTGAAGGCTGGGAGGAGGACGCGCACG 407
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThSera 134
408 ACAGATTCACTCCACTACGAGAGCGCGCTTGAGCATCACACGCTG 457
134 sPaRgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
458 ACCGTGACCGCTAATAGTATGTTGTGGCGCGCTAGCTGTGGAAGCC 507
151 GlyPheAspTrrValTyrTyrGluSerArgAsnHisValHisValSera 167
508 GGATTGACAGTGGGTCTACTACGAGTCCCGCACACATCCACGATG 557
167 llySAlaAspAsnSerLeuAlaValArgAlaGlyGly 179
558 CAAAGCTGATTACTCACTGCGGTCCGAGCGGAGGC 594

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:us-08-460-900C-2

seq_documentation_block:
; Sequence 2, Application US/08460900C
; Patent No. 6165747
; GENERAL INFORMATION:
```

APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
APPLICANT: Bumcrot, David A.  
APPLICANT: Marti-Gorostiza, Elisa  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,900C  
FILING DATE: 5-JUNE-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/435,093  
FILING DATE: 4-MAY-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/356,060  
FILING DATE: 14-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/176,427  
FILING DATE: 30-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006.05  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 832-1000  
TELEFAX: (617) 832-7000  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1190 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1188  
US-08-460-900C-2

alignment\_scores:  
Quality: 932.00 Length: 179  
Ratio: 5.207 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 97.207

alignment\_block:  
US-08-900-220C-17\_COPY\_20\_198 x US-08-460-900C-2 ..  
Align seg 1/1 to: US-08-460-900C-2 from: 1 to: 1190

1 AAgInSerCysGlyProGlyArgGlyProValGlyArgArgTyrAl 17  
|||||  
58 CCGCAGACTGGGGCGGGCGCGAGACCGGTGGCCGCGCGCTATGCT 107  
17 AArgGlnLeuValProLeuLeuTyrGlnPheValProGlyValP 34  
|||||  
108 GCGCAGCAACTGTGCTGCTGTACAGACAGTTGTGCGCCACTATGC 157  
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50  
|||||  
158 CCGAGCGGACCTGGGGCGGAGTGGCCAGCGGAGGAGGAGGTAAACAAG 207

51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIle1 67  
|||||  
208 GCGTCGAGCGCTTCGGGAGCTCGTACCCACTACACACCCGACATAT 257  
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84  
|||||  
258 CTTCAAGGATGAGGAGAACAGCGCGCAGACCGCTGTATGACAGAGCTT 307  
84 ySLysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100  
|||||  
308 GCAAGAGCGGCTGAACGCTTACGCTATCGCGGTATGAACATGTGGCC 357  
101 GlyValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAla 117  
|||||  
358 GGAGTACCGCTACGCTGCTACGCTGAAGCGTGGACGACGACGACACGC 407  
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSerA 134  
|||||  
408 ACAGGATTCACCTCCACTACGAGAGCGCGCTTGACATCACACAGCTCG 457  
134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150  
|||||  
458 ACCGTGACCGGTATATAGTATGTGTGTGGCGGCTAGCTGTGAAGCC 507  
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVa 167  
|||||  
508 GGATTGCACTGGCTCTACAGACAGTCCGCCACACCATCCACTATCGCT 557  
167 lLysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179  
|||||  
558 CAAAGCTGATTAACCTGCGCGGTCCGAGCCGAGGCG 594

seq\_name: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:us-08-176-427B-1

seq\_documentation\_block:  
Sequence 1, Application US/08176427B  
Patent No. 5789543  
GENERAL INFORMATION:  
APPLICANT: Ingham, Phillip W.  
APPLICANT: McMahon, Andrew P.  
APPLICANT: Tablin, Clifford J.  
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
TITLE OF INVENTION: Proteins and Uses Related Thereto  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/176,427B  
FILING DATE: 30-DEC-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: HMV-006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1277 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both

TOPOLOGY: linear  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..1277  
 US-08-176-427B-1

alignment\_scores:  
 Quality: 741.00 Length: 177  
 Ratio: 4.411 Gaps: 2  
 Percent Similarity: 94.915 Percent Identity: 74.011

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x US-08-176-427B-1 ..

Align seg 1/1 to: US-08-176-427B-1 from: 1 to: 1277

```

3  SerCysGlyProGlyArgGlyProValGlyArgArgGlyValArgGly 19
   :::::::::::::::::::::::::::::
76  ACTGTGTGACCAAGGAGGAGG...ATTGMAAAAGAGGACCCCAAAA 122
   :::::::::::::::::::::::::::::
19  sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGly 36
   | :::::::::::::::::::::
123  G...CTGACCCCGTTAGCCTATAGCAGTTTATTCCTCATGTGGCAGAGA 169
   :::::::::::::::::::::::::::::
36  rGThrLeuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
   :::::::::::::::::::::::::::::
170  AGACCCCTAGGAGGCGCAGTGAAGATGTGAAGGAGATCACAGAACTCC 219
   :::::::::::::::::::::::::::::
53  GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
   :::::::::::::::::::::::::::::
220  GAGGATTTAAAGAACTAACCCCAATTACACCCCTGACATTTATTTTAA 269
   :::::::::::::::::::::::::::::
69  sAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgCysLys 86
   :::::::::::::::::::::::::::::
270  GGATGAGAGACACGCGACCTGCACATGATGACTACGCGCTGCAGAG 319
   :::::::::::::::::::::::::::::
86  LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
   :::::::::::::::::::::::::::::
320  ACAAGCTGAATGCCCTGGCGATCTCGGTGATGAACAGTGCGCGGGTGG 369
   :::::::::::::::::::::::::::::
103  ArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHisAlaGlnAs 119
   :::::::::::::::::::::::::::::
370  AAGCTGCGGCTGACCGAGGCTGGGAGAGAGATGCCATCTCCGAGGA 419
   :::::::::::::::::::::::::::::
119  pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
   :::::::::::::::::::::::::::::
420  ATGCTCTCATTAGSAGGCTGCCGCCCTGCGACATCCACAGCTCGATGCG 469
   :::::::::::::::::::::::::::::
136  sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
   :::::::::::::::::::::::::::::
470  ACCGACACAGAGTACGGAAATGCTGCCCGCTCGCGAGCGCGGCTTC 519
   :::::::::::::::::::::::::::::
153  AspThrValTyrTyrGlySerArgAsnHisValHisValSerValLysAl 169
   :::::::::::::::::::::::::::::
520  GACTGGTCTTACTACGAGTCCAGAGCGCAATCATCAGTCTCCGTAAGC 569
   :::::::::::::::::::::::::::::
169  aAspAsnSerLeuAlaValArgAlaGlyLys 179
   :::::::::::::::::::::::::::::
570  AGAAACTCAAGTGCACGAAATGCAGAGG 600
   :::::::::::::::::::::::::::::

```

seq\_name: /cgn2\_6/plodata/2/lna/5B\_COMB.seq:US-08-356-060A-1

seq\_documentation\_block:

; Sequence 1, Application US/08356060A  
 ; Patent No. 5844079  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ingham, Phillip W.  
 ; APPLICANT: McMahon, Andrew P.  
 ; APPLICANT: Tablin, Clifford J.  
 ; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing  
 ; TITLE OF INVENTION: Proteins and Uses Related Thereto

```

; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,060A
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMT-006CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1277 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1275
; US-08-356-060A-1

```

alignment\_scores:  
 Quality: 741.00 Length: 177  
 Ratio: 4.411 Gaps: 2  
 Percent Similarity: 94.915 Percent Identity: 74.011

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x US-08-356-060A-1 ..

Align seg 1/1 to: US-08-356-060A-1 from: 1 to: 1277

```

3  SerCysGlyProGlyArgGlyProValGlyArgArgGlyValArgGly 19
   :::::::::::::::::::::::::::::
76  ACTGTGTGACCAAGGAGGAGG...ATTGMAAAAGAGGACCCCAAAA 122
   :::::::::::::::::::::::::::::
19  sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGly 36
   | :::::::::::::::::::::
123  G...CTGACCCCGTTAGCCTATAGCAGTTTATTCCTCATGTGGCAGAGA 169
   :::::::::::::::::::::::::::::
36  rGThrLeuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
   :::::::::::::::::::::::::::::
170  AGACCCCTAGGAGGCGCAGTGAAGATGTGAAGGAGATCACAGAACTCC 219
   :::::::::::::::::::::::::::::
53  GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
   :::::::::::::::::::::::::::::
220  GAGGATTTAAAGAACTAACCCCAATTACACCCCTGACATTTATTTTAA 269
   :::::::::::::::::::::::::::::
69  sAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgCysLys 86
   :::::::::::::::::::::::::::::
270  GGATGAGAGACACGCGGAGCTGACAGATGATGACTACGCGCTGCAGAG 319
   :::::::::::::::::::::::::::::
86  LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
   :::::::::::::::::::::::::::::
320  ACAAGCTGAATGCCCTGGCGATCTCGGTGATGAACAGTGCGCGGGTGG 369
   :::::::::::::::::::::::::::::

```

```

103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AAGCTGGCGGTACCGAGCGCTGGGACGACGATGGCCATCATCTCCGAGGA 419
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgA 136
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATCGCTGACACAGAGGCTCGCGCGGTGACATCCACGATCGGATCGGG 469
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCGCGAGCAATACGGAATGCTGGCCCGCTCGCGGTGAGCGCGCTTC 519
153 AsPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GACGTGGCTCTACTACGACTCCAGCGCGCACATCCATCTGCTCGTCAAAAGC 569
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 AGAAAACTCAGTGGCGAGCAATCAGAGGCG 600

seq_name: /cgn2_6/ptodata/2/lna/5b_COMB.seq:us-08-460-900c-1

seq_documentation_block:
: Sequence 1, Application US/08460900C
: Patent No. 6165747
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tablin, Clifford J.
: APPLICANT: Bumcrot, David A.
: APPLICANT: Marti-Gorostiza, Elisa
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY, HOAG & ELIOT LLP
: STREET: One Post Office Square
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,900C
: FILING DATE: 5-JUNE-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/435,093
: FILING DATE: 4-MAY-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/356,060
: FILING DATE: 14-DEC-1994
: APPLICATION DATA:
: APPLICATION NUMBER: US 08/176,427
: FILING DATE: 30-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: HMV-006.05
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 832-1000
: TELEFAX: (617) 832-7000
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1277 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA

```

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1275
: US-08-460-900c-1

alignment_scores:
  Quality: 741.00      Length: 177
  Ratio: 4.411        Gaps: 2
  Percent Similarity: 94.915   Percent Identity: 74.011

alignment_block:
  us-08-900-220c-17_copy_20_198 x us-08-460-900c-1 ..
  Align seg 1/1 to: us-08-460-900c-1 from: 1 to: 1277

19 scLnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGluA 36
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 G...CTGACCCCGTTAGCCTATACGACGTTATTCTCCAAATGTGGCAGAGA 169
36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgLysSer 52
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
170 AGACCCCTAGGGCGGCAGTGAAGATATCAAGCGACATCCACAGAAACTCC 219
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLys 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
220 GAGAGATTAAAGAACTAACCCCAATTACACCTGCATATTATTTTAA 269
69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 GGATGACAGACACGCGGAGCTGACAGACTGATGACTCAGCCCTCAGAG 319
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
320 ACAAGCTGAATGCCCTGGGAGATCTCGTGATGATACACAGATGGCCGGGG 369
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAlaGlnAs 119
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
370 AAGCTGGCGGTACCGAGGCTGGGACGAGATGGCCATCTCCGAGGA 419
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgA 136
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
420 ATCGCTGACACAGAGGCTCGCGCGGTGACATCCACGATCGGATCGGG 469
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
470 ACCGCGAGCAATACGGAATGCTGGCCCGCTCGCGGTGAGCGCGCTTC 519
153 AsPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 GACTGGGTCTACTACGAGTCCAGGCGCGCACATCCATCTGCTCGTCAAAAGC 569
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
570 AGAAAACTCAGTGGCGAGCAATCAGAGGCG 600

seq_name: /cgn2_6/ptodata/2/lna/5a_COMB.seq:us-08-176-427b-7

seq_documentation_block:
: Sequence 7, Application US/08176427B
: Patent No. 5789543
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tablin, Clifford J.
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,427B
FILING DATE: 30-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMT-006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1313 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1314
US-08-176-427B-7

```

```

alignment_scores:
Quality: 740.00      Length: 177
Ratio: 4.431         Gaps: 2
Percent Similarity: 94.350   Percent Identity: 74.576

```

```

alignment_block:
US-08-900-220C-17_COPY_20_198 x US-08-176-427B-7

```

```

Align seg 1/1 to: US-08-176-427B-7 from: 1 to: 1313

```

```

3 SerGcGlyProGlyArgGlyProValGlyArgArgGlyValArgGly 19
:::|||||:::|||||:::|||||:::|||||:::|||||
70 GCCTGTGGCCGCGCAGGAGG...TTTGAAGAGCGCGCACCCCAAAA 116
19 scGlnLeuValProLeuLeuTyrGlnPheValProGlyValProGln 36
| |||:::||||| |||:::|||||:::|||||:::|||||
117 G...CTGACCCCTTACCTACACAGCAGTTTATCCCAAGGTAGCCGAGA 163
36 rGlnrLeuGlyAlaSerGlyProAlaGlnGlyArgValAlaArgGlySer 52
:::|||||:::|||||:::|||||:::|||||:::|||||
164 AGACCCAGGAGGCGCAGCGCATATGAGGAAGATCACACAGAAATCC 213
53 GlnArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
|||||:::|||||:::|||||:::|||||:::|||||
214 GAACGATTTAAGAACTCACCCCAATTACACCCGACATCATATTAA 263
69 sAspGlnGlnAsnSerGlyAlaAspArgLeuMetThrGlnArgCysArg 86
|||||:::|||||:::|||||:::|||||:::|||||
264 GCGATGAGAAACACGCGACACGCGGTGATGACTCAGAGGCGCAAG 313
86 IuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTyrProGlyVal 102
:::|||||:::|||||:::|||||:::|||||:::|||||
314 ACAAGTTAAATGCTTGCCATCTCTGTGATGAAACAGTGCCTGGAGTG 363
103 ArgLeuArgValThrGlnGlyTyrPaspGlnAspGlyIshisAlaGlnAs 119
|||||:::|||||:::|||||:::|||||:::|||||
364 AGCGTCGAGTAGACGAGGCGGTGATGAGACGCGCATCATTCAGAGGA 413

```

```

119 pSerLeuHisTyrGlnGlyArgAlaLeuAspIleThrThrSerAspArg 136
:|||||:::|||||:::|||||:::|||||:::|||||
414 GTCTTAACCTATGAGGTCGACGACATCCACCGACGTCGACCGGG 463
136 sPArgAsnTyrTyrGlyLeuLeuAlaArgLeuAlaValGlnAlaGlyPhe 152
|||||:::|||||:::|||||:::|||||:::|||||
464 ACCGACAGATGACGCGATCTGCTCGCTGCTGCTGAGACGAGTTTC 513
514 GACTGGGTACTATGATGAACTCAAGCTCACATCCACTGCTGTCTGAAGC 563
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|:::|||||:::|||||:::|||||
564 AGAGAACTCCGTGGCGCCAAATCCGCGCGC 594

```

```

seq_name: /gcn2_6/ptodata/2/lna/5B_COMB.seq:US-08-356-060A-4

```

```

seq_documentation_block:

```

```

Sequence 4, Application US/08356060A
Patent No. 5844079

```

```

GENERAL INFORMATION:

```

```

APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tablin, Clifford J.

```

```

TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing

```

```

NUMBER OF SEQUENCES: 47

```

```

CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: LAHIVE & COCKFIELD

```

```

STREET: 60 State Street

```

```

CITY: Boston

```

```

STATE: MA

```

```

COUNTRY: USA

```

```

ZIP: 02109

```

```

COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: ASCII(text)

```

```

CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/356,060A

```

```

FILING DATE: 14-DEC-1994

```

```

CLASSIFICATION: 435

```

```

PRIORITY APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/176,427

```

```

FILING DATE: 30-DEC-1993

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Vincent, Matthew P.

```

```

REGISTRATION NUMBER: 36,709

```

```

REFERENCE/DOCKET NUMBER: HMT-006CP

```

```

TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (617) 227-7400

```

```

TELEFAX: (617) 227-5941

```

```

INFORMATION FOR SEQ ID NO: 4:

```

```

SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 1313 base pairs

```

```

TYPE: nucleic acid

```

```

STRANDEDNESS: both

```

```

TOPOLOGY: linear

```

```

MOLECULE TYPE: cDNA

```

```

FEATURE:

```

```

NAME/KEY: CDS

```

```

LOCATION: 1..1314

```

```

US-08-356-060A-4

```

```

alignment_scores:
Quality: 740.00      Length: 177
Ratio: 4.431         Gaps: 2
Percent Similarity: 94.350   Percent Identity: 74.576

```

```

alignment_block:

```

US-08-900-220C-17\_COPY\_20\_198 x US-08-356-060A-4 ..

Align seq 1/1 to: US-08-356-060A-4 from: 1 to: 1313

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgGlyArgGly 19
  ::::::::::::::::::::
70 GCCGTGGGCCCCGAGAGGGG...TTTGGAAAGAGGGGACCCCAAAA 116
  ::::::::::::::::::::
19 SGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGln 36
  ::::::::::::::::::::
117 G...CTGACCCCTTACCTAGCAGAGGATTTATTCACAGCTAGCCGAGA 163
  ::::::::::::::::::::
36 rgtHLeuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
  ::::::::::::::::::::
164 AGACCCAGAGGGCCGAGGAGATGAGAGGAGATCACAAGAAAGCTCC 213
  ::::::::::::::::::::
53 GluArgPheArgAspLeuValProAsnTyrTrpAsnProAspIleIlePhe 69
  ::::::::::::::::::::
214 GAGCGATTTAAGAACTACCCCAATTACAAACCCGACATCATATTAA 263
  ::::::::::::::::::::
69 SASpGluLysSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  ::::::::::::::::::::
264 GGATGAGAGAAACACGGAGGAGCAGCGGCTGATGATCAGAGGTCAAG 313
  ::::::::::::::::::::
86 IuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  ::::::::::::::::::::
314 ACAAGTTAAATGCTTGGCCATCTGTGATGAACCAAGTCCCTGAGTG 363
  ::::::::::::::::::::
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAlaGlnAs 119
  ::::::::::::::::::::
364 AGGCTGGAGAGGACCGAGGCTGGGATGAGAGGCCATCTTCAGAGGA 413
  ::::::::::::::::::::
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTrpSerAspArg 136
  ::::::::::::::::::::
414 GTCTCTACATATGAGGATGAGGATGAGATGACACACGTCACAGCCGG 463
  ::::::::::::::::::::
136 SPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
  ::::::::::::::::::::
464 ACCGACCAAGTACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
  ::::::::::::::::::::
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
  ::::::::::::::::::::
514 GACGGGCTTCTATGATTCACAAAGCTCACAATCCATCTTCTGTGAAGC 563
  ::::::::::::::::::::
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
  ::::::::::::::::::::
564 AGAGAACTCCGTGGCGGCAAAATCCGCGCGC 594
  ::::::::::::::::::::

```

seq\_name: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:US-08-460-900C-4

seq\_documentation\_block:

```

: Sequence 4, Application US/08460900C
: Patent No. 6155747
: GENERAL INFORMATION:
: APPLICANT: Ingham, Phillip W.
: APPLICANT: McMahon, Andrew P.
: APPLICANT: Tabin, Clifford J.
: APPLICANT: Bumcrot, David A.
: APPLICANT: Marti-Gorostiza, Elisa
: TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
: TITLE OF INVENTION: Proteins and Uses Related Thereto
: NUMBER OF SEQUENCES: 62
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY, HOAG & ELIOT LLP
: STREET: One Post Office Square
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/460,900C
: FILING DATE: 5-JUNE-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/435,093
: FILING DATE: 4-MAY-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/356,060
: FILING DATE: 14-DEC-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/176,427
: FILING DATE: 30-DEC-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Vincent, Matthew P.
: REGISTRATION NUMBER: 36,709
: REFERENCE/DOCKET NUMBER: HMV-006.05
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 832-1000
: TELEFAX: (617) 832-7000
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1313 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..1311
: US-08-460-900C-4

```

```

alignment_scores:
  Quality: 740.00      Length: 177
  Ratio: 4.431        Gaps: 2
  Percent Similarity: 94.350      Percent Identity: 74.576

```

alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x US-08-460-900C-4 ..

Align seq 1/1 to: US-08-460-900C-4 from: 1 to: 1313

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgGlyArgGly 19
  ::::::::::::::::::::
70 GCCGTGGGCCCCGAGAGGGG...TTTGGAAAGAGGGGACCCCAAAA 116
  ::::::::::::::::::::
19 SGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGln 36
  ::::::::::::::::::::
117 G...CTGACCCCTTACCTAGCAGAGGATTTATTCACAGCTAGCCGAGA 163
  ::::::::::::::::::::
36 rgtHLeuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
  ::::::::::::::::::::
164 AGACCCAGAGGGCCGAGGAGATGAGAGGAGATCACAAGAAAGCTCC 213
  ::::::::::::::::::::
53 GluArgPheArgAspLeuValProAsnTyrTrpAsnProAspIleIlePhe 69
  ::::::::::::::::::::
214 GAGCGATTTAAGAACTACCCCAATTACAAACCCGACATCATATTAA 263
  ::::::::::::::::::::
69 SASpGluLysSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
  ::::::::::::::::::::
264 GGATGAGAGAAACACGGAGGAGCAGCGGCTGATGATCAGAGGTCAAG 313
  ::::::::::::::::::::
86 IuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
  ::::::::::::::::::::
314 ACAAGTTAAATGCTTGGCCATCTGTGATGAACCAAGTCCCTGAGTG 363
  ::::::::::::::::::::
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAlaGlnAs 119
  ::::::::::::::::::::
364 AGGCTGGAGAGGACCGAGGCTGGGATGAGAGGCCATCTTCAGAGGA 413
  ::::::::::::::::::::
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTrpSerAspArg 136
  ::::::::::::::::::::

```





```

APPLICATION NUMBER: US/08/460,900C
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 4-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1425
US-08-460-900C-6

```

```

alignment_scores:
Quality: 738.00      Length: 177
Ratio: 4.419         Gaps: 2
Percent Similarity: 94.350      Percent Identity: 74.011

```

```

alignment_block:
US-08-900-220c-17_copy_20_198 x US-08-460-900C-6 ..

```

```

Align seg 1/1 to: US-08-460-900C-6 from: 1 to: 1425

```

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAlaArgly 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 GGGTGGCGGACCGGCGGAGGGG...TTCCGGAGAGAGAGAGCCCAAAA 113
19 scInleuValProleuLeuTyrLysGlnPheValProGlyValProGlnA 36
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
114 G...CTGACCCCTTTAGCCTACAGCAGTTTATCCCAATGTGGCCGAGA 160
36 rGthrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArglySer 52
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
161 AGACCTAGGCGCCAGCGAGAGGTATGAAGGAGAGATCTCCAGAAATC 210
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleLeuPhe 69
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
211 GAGCGATTTTAAAGAACTCACCCCAATTACAAACCCGACATCATATTAA 260
69 sAspGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
261 GGATGAGAAACACCGGAGCGGACAGCTGATGACTCAGAGGTGAAG 310
86 luArgValAsnAlaLeuAlaIleAlaValMetAsnMetTyrProGlyVal 102
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
311 ACAAGTTGAACGTTTGGCATCTCGGTATGATACACAGTGGCAGAGATG 360
103 ArgLeuArgValThrGluGlyTyrAspGluAspArgLysHisAlaGln 119
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
361 AAACCTGGGGTAGACGAGGGCTGGGACGAAGATGGCCACCTCAAGGA 410
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
411 GTCTGTGACCTAGAGAGCGCGGCACTGACATCACACAGCTGTGACCGG 460

```

```

136 sPArGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
461 ACCGAGCAAGTAGCGCATGTGGCCCGCTGGCGGTGAGCGCGGCTTC 510
153 AspTyrValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
511 GACTGGGTGTACTACGAGTCCAGGCGACATATTCACCTGCTCGTGAAGC 560
169 asPAseSerLeuAlaValArgAlaGlyLys 179
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
561 AGAGAACTGCTGGCGCGCCCAATCGGGAGGC 591

```

```

seq_name: /cgn2_6/prodata/2/ina/5A_COMB.seq:us-08-748-591-10

```

```

seq_documentation_block:

```

```

Sequence 10, Application US/08748591

```

```

Patent No. 5759811

```

```

GENERAL INFORMATION:

```

```

APPLICANT: Epstein, Ervin

```

```

APPLICANT: Hu, Zhilan

```

```

APPLICANT: Bonifas, Jeanette

```

```

TITLE OF INVENTION: Mutant Human Hedgehog Gene

```

```

NUMBER OF SEQUENCES: 23

```

```

CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Fish and Richardson

```

```

STREET: 2200 Sand Hill Road

```

```

CITY: Menlo Park

```

```

STATE: CA

```

```

COUNTRY: USA

```

```

ZIP: 94025

```

```

COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/08/748,591

```

```

FILING DATE:

```

```

CLASSIFICATION: 435

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Sherwood, Pamela J

```

```

REGISTRATION NUMBER: 36,677

```

```

REFERENCE/DOCKET NUMBER: 06510/067001

```

```

TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (415) 322-5070

```

```

TELEFAX: (415) 854-0875

```

```

INFORMATION FOR SEQ ID NO: 10:

```

```

SEQUENCE CHARACTERISTICS:

```

```

LENGTH: 1576 base pairs

```

```

TYPE: nucleic acid

```

```

STRANDEDNESS: single

```

```

TOPOLOGY: linear

```

```

MOLECULE TYPE: cDNA

```

```

US-08-748-591-10

```

```

alignment_scores:

```

```

Quality: 734.00      Length: 177

```

```

Ratio: 4.395         Gaps: 2

```

```

Percent Similarity: 94.350      Percent Identity: 73.446

```

```

alignment_block:
US-08-900-220c-17_copy_20_198 x US-08-748-591-10 ..

```

```

Align seg 1/1 to: US-08-748-591-10 from: 1 to: 1576

```

```

3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAlaArgly 19
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 GCGTGGGAGACCGGCGGAGGGG...TTCCGGAGAGAGAGAGCCCAAAA 264
19 scInleuValProleuLeuTyrLysGlnPheValProGlyValProGlnA 36
| |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
265 G...CTGACCCCTTTAGCCTACAGCAGTTTATCCCAATGTGGCCGAGA 311

```

```

36  rgrhrleuglyalaSerGlyProAlaGluGlyAlaArgGlySer 52
312 AGACCCGAGGGCCAGCGAGGATGAGAGGAGGATCTCCAGAACTCC 361
53  GluArgPheArgPheValProAsnTyrAsnProAspIleIlePhe 69
362 GAGCGATTTAAGAACTACCCCAATTACACCCCGACATCATATTTAA 411
69  sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysIysG 86
412 GGATGAGAAACACGAGCGGACGAGCTGATGATCAGAGGTGTAAG 461
86  LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
462 ACAAGTTGAACGCTTTGGCCATCTCGGTGATAAACAGTGGCCGAGGTG 511
103 ArgLeuArgValThrGluGlyTyrPAspGluAspGlyHisHisAlaGlnAs 119
512 AAATCGCGGTGACCGAGGCTGGAGGAGAGATGCCACCACTAGAGA 561
119 pSerLeuHisTyrGluGlyAlaAlaLeuAspIleThrThrSerAspArg 136
562 GTCTCTGACTACGAGGCGCGCGAGTGCATCACACGCTGACCGCG 611
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
612 ACCCCACCAAGTAGCGCATGCTGGCCCGCTGGCGGTGAGGCGCGCTTC 661
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
662 GACGGGTGTACTACGAGTCCAGAGGACATATTCACATGCTCGTGAAAGC 711
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
712 AGAGACTCGGTGGCGCCAAATGCGAGGC 742

```

seq\_name: /cgn2\_6/ptodata/2/1na/5A\_COMB.seq:US-08-748-591-5

```

seq_documentation_block:
; Sequence 5, Application US/08748591
; Patent No. 5759811
; GENERAL INFORMATION:
; APPLICANT: Epstein, Ervin
; APPLICANT: Hu, Zhilan
; APPLICANT: Bonifas, Jeanette
; TITLE OF INVENTION: Mutant Human Hedgehog Gene
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish and Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,591
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: 06510/067001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 322-5070
; TELEFAX: (415) 854-0875
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 1576 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-748-591-5

```

```

alignment_scores:
  Quality: 732.00      Length: 177
  Ratio: 4.383        Gaps: 2
  Percent Similarity: 94.350  Percent Identity: 73.446

```

alignment\_block:

US-08-900-220c-17\_copy\_20\_198 x US-08-748-591-5 ..

Align seg 1/1 to: US-08-748-591-5 from: 1 to: 1576

```

3  SerCysGlyProGlyAlaArgGlyProValGlyArgArgGlyAlaArgGly 19
218 GCGTGGCGGACCGGCGAGGGG...TTGCGGAGAGAGGAGGACCCCAAAA 264
19  sGluLeuValProLeuLeuTyrLysGlnPheValProGlyValProGlu 36
265 G...CTACCCCTTACCTACAGCAAGTTATCCCAATGTGGCGGAGA 311
36  rgrhrleuglyalaSerGlyProAlaGluGlyAlaArgGlySer 52
312 AGACCCGAGGGCCAGCGAGGATGAGAGGAGGATCTCCAGAACTCC 361
53  GluArgPheArgPheValProAsnTyrAsnProAspIleIlePhe 69
362 GAGCGATTTAAGAACTACCCCAATTACACCCCGACATCATATTTAA 411
69  sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysIysG 86
412 GGATGAGAAACACGAGCGGACGAGCTGATGATCAGAGGTGTAAG 461
86  LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetThrProGlyVal 102
462 ACAAGTTGAACGCTTTGGCCATCTCGGTGATAAACAGTGGCCGAGGTG 511
103 ArgLeuArgValThrGluGlyTyrPAspGluAspGlyHisHisAlaGlnAs 119
512 AAATCGCGGTGACCGAGGCTGGAGGAGAGATGCCACCACTAGAGA 561
119 pSerLeuHisTyrGluGlyAlaAlaLeuAspIleThrThrSerAspArg 136
562 GTCTCTGACTACGAGGCGCGCGAGTGCATCACACGCTGACCGCG 611
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
612 ACCCCACCAAGTAGCGCATGCTGGCCCGCTGGCGGTGAGGCGCGCTTC 661
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
662 GACGGGTGTACTACGAGTCCAGAGGACATATTCACATGCTCGTGAAAGC 711
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
712 AGAGACTCGGTGGCGCCAAATGCGAGGC 742

```

seq\_name: /cgn2\_6/ptodata/2/1na/6B\_COMB.seq:US-08-460-900C-3

```

seq_documentation_block:
; Sequence 3, Application US/08460900C
; Patent No. 6185747
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tablin, Clifford J.
; APPLICANT: Bumcrot, David A.
; APPLICANT: Marti-Gorostiza, Elisa

```



US-08-176-427B-9

alignment\_scores:                   Quality: 723.00                   Length: 177  
                                      Ratio: 4.409                   Gaps: 2  
Percent Similarity: 92.655           Percent Identity: 72.881

alignment\_block:

US-08-900-220c-17\_copy\_20\_198 x US-08-176-427B-9 ..

Align seg 1/1 to: US-08-176-427B-9 from: 1 to: 1256

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgGlyArgGly 19
:::|||||:::|||||:::|||||:::|||||:::|||||
67 GCCGCGGCTCTGGCAGAGC...TACGCCAGAGAGACATCCGAGAA 113
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGly 36
|:::|||||:::|||||:::|||||:::|||||:::|||||
114 G...CTGCACACTCTCGCCCTACAGCAGTTCATCTAATCTCCCGAGA 160
36 rGThrLeuGlyAlaSerGlyProAlaGlnGlyArgValAlaArgGlySer 52
:::|||||:::|||||:::|||||:::|||||:::|||||
161 AGACCTTAGGGCCAGCGGCAGATACAGGCGAAGATACGCCCATTCG 210
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
|||||:::|||||:::|||||:::|||||:::|||||
211 GAGGATTTAAGAACTTACTCCAAATTACAAATCCGACATTAATCTTAA 260
69 sAspGlnGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
|||||:::|||||:::|||||:::|||||:::|||||
261 GGATGAGAGAAACCGGAGCGGACGAGCTCATGACACAGATGCAAG 310
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
:::|||||:::|||||:::|||||:::|||||:::|||||
311 ACAAGCTGAAGCTGCGCATCTCTAATGAACCACTGCGCGGCGTT 360
103 ArgLeuArgValThrGlnGlyTrpAspGluAspGlyHisHisAlaGlnAs 119
:::|||||:::|||||:::|||||:::|||||:::|||||
361 AAGCTGCTGTGACAGAGGCTGGGATGAGGACGCTCACCATTTTGAAGA 410
119 pSerLeuHisTyrGlnGlyArgAlaLeuAspIleThrThrSerAspArg 136
:::|||||:::|||||:::|||||:::|||||:::|||||
411 ATCACTCCACTACAGAGGAGAGCTGTTCATATTACCACTCTGACCGAG 460
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|:::|||||:::|||||:::|||||:::|||||:::|||||
461 ACAAGACCAATACGGGACACTGTCTGCTGAGCTGAGGCTGATTT 510
153 AspTrpValTyrTyrGlnSerArgAsnHisValHisValSerValLysAl 169
|||||:::|||||:::|||||:::|||||:::|||||
511 GACTGGGTCTATTACGAGTCCAAAGCCACATTCATTCCTGTGCAAGC 560
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|:::|||||:::|||||:::|||||:::|||||
561 AGAAATTCGTTGCTGCGAAATCTGCGGGC 591
```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 22, 2001, 11:01:20 ; Search time 57.94 Seconds

(without alignments)  
362.102 Million cell updates/sec

Title: US-08-900-220c-17\_COPY\_20\_198

Perfect score: 950  
Sequence: 1 AOSCGPGGPGVGRRRYARKQ.....RNHHVSVKADNSLAVRAG 179

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL\_15:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	716	75.4	414	13	09W709 paralicthys
2	715	75.3	406	13	057567 notophthalm
3	712	74.9	434	13	057404 pleurodeles
4	664	69.9	129	11	09WUP6 09wup6 ratus norv
5	640	67.4	415	5	017499 017499 branchiosto
6	634	66.7	442	13	073803 073803 fugu rudrip
7	618	65.1	139	6	09XSI6 09xsi6 bos taurus
8	600.5	63.2	471	5	09VQC4 09vqc4 drosophila
9	599	63.1	138	13	09WEC1 09wec1 eleutheroda
10	563.5	59.3	410	5	061676 061676 lytechinus
11	550	57.9	185	5	096699 096699 junonia coe
12	525	55.3	150	13	09YGV7 09ygv7 ambystoma m
13	523	55.1	161	11	09RI79 09ri79 ratus norv
14	457	48.1	177	11	09WV29 09wv29 ratius norv
15	375	39.5	88	13	09YU33 09yu33 brachydanio
16	367	38.6	87	5	09TX30 09tx30 anopheles g
17	352	37.1	80	13	042441 042441 oryzias lat
18	340	35.8	227	5	09U5Z6 09u5z6 branchiosto
19	331	34.8	119	13	042128 042128 oryzias lat

## ALIGNMENTS

20	310	32.6	63	13	09IB14	09ib14 rana catesb
21	305	32.1	64	13	09PRF5	09prf5 oryzias lat
22	228	24.0	49	5	09TX33	09tx33 hirudo medi
23	219	23.1	48	5	09TX31	09tx31 tiribolium c
24	171.5	18.1	48	5	09TX32	09tx32 strongyloce
25	90	9.5	127	2	09X758	09x758 streptomyce
26	89.5	9.4	490	8	P97956	P97956 unidentified
27	86	9.1	54	13	042233	042233 coturnix co
28	85.5	9.0	486	4	09UNFO	09unfo homo sapien
29	85	8.9	484	4	09LAV5	09lav5 thermomono
30	84.5	8.9	460	8	078521	078521 botrydium s
31	83.5	8.8	490	8	098947	098947 thalassios
32	81.5	8.6	342	14	058622	058622 royal farm
33	80.5	8.5	490	8	098945	098945 detonula co
34	79.5	8.4	471	5	023488	023488 caenorhabdi
35	78.5	8.3	476	8	046962	046962 rhizosoleni
36	78.5	8.3	680	2	052747	052747 ruminococcu
37	78.5	8.3	1193	4	092580	092580 homo sapien
38	78	8.2	675	10	09MX05	09mx05 arabidopsis
39	77.5	8.2	330	2	008347	008347 streptomyce
40	77.5	8.2	474	8	047285	047285 bumilleriop
41	77.5	8.2	488	8	09XPY4	09xpy4 lammaricoc
42	77.5	8.2	523	14	09QPR1	09qpr1 canine dist
43	77.5	8.2	3670	2	09Z4X5	09z4x5 streptomyce
44	77	8.1	1124	5	09U2C0	09u2c0 caenorhabdi
45	76.5	8.1	377	8	046979	046979 mallomonas

RESULT 1  
ID 09W709 PRELIMINARY: PRT: 414 AA.  
AC 09W709;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE SONIC HEDGEHOG.  
GN SHH.  
OS Paralicthys olivaceus (Flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Bothidae; Paralicthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99238226; PubMed=10223710;  
RA Suzuki T., Ichiro O., Kurokawa T.;  
RT "Retinoic acid given at late embryonic stage depresses sonic hedgehog  
RT and Hoxd-4 expression in the pharyngeal area and induces skeletal  
RT malformation in flounder (Paralicthys olivaceus) embryos.";  
RL Dev. Growth Differ. 41:143-152(1999).  
DR EMBL: AB029748; BAA82360.1; -.  
DR HSSP: Q62226; IVH.  
DR INTERPRO: IPR000320; -.  
DR INTERPRO: IPR001657; -.  
DR INTERPRO: IPR001677; -.  
DR PFM: PF01079; Hint: 1.  
DR PFM: PF01085; HH\_signal: 1.  
DR PRINTS: PR00632; SONICHHOG.  
SQ SEQUENCE 414 AA: 45945 MM: 50607/BF3DB7C0DA3 CRC64;

Query Match 75.4%; Score 716; DB 13; Length 414;  
Best Local Similarity 72.7%; Pred. No. 2,7e-58;  
Matches 128; Conservative 28; Mismatches 18; Indels 2; Gaps 2;

Oy 4 CGPGGPGVGRRRYARKQVPLLYKQFVGVEERTIGASGPAEGRVRSERFDLYPVYN 63  
Db 24 CGPGGPGVGRRRYARKQVPLLYKQFVGVEERTIGASGPAEGRVRSERFDLYPVYN 81

QY 64 PDIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 123  
 DB 82 TDIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 141  
 QY 124 EGRALDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 179  
 DB 142 EGRAVDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 197

## RESULT 2

ID 057567 PRELIMINARY: PRT: 406 AA.

AC 057567  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE HEDGHO SEGMENT POLARITY HOMOLOG.  
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;  
 OC Notophthalmus.  
 NCBI\_TaxID=8316;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Stark D.R., Gates P.B., Brookes J.P., Ferretti P.  
 RL Dev. Dyn. 0:0-0(1998).  
 DB EMBL: AF047466; AAC03108.1; -.  
 DR HSSP: O62226; 1VHH.  
 DR INTERPRO: IPR000320; -.  
 DR INTERPRO: IPR001657; -.  
 DR INTERPRO: IPR001767; -.  
 DR INTERPRO: IPR002375; -.  
 DR PFAM: PF01079; HIntc. 1.  
 DR PFAM: PF01085; HH\_Signal; 1.  
 DR PRINTS: PR00632; SONICHHOG.  
 DR PROSITE: PS00103; PTR\_PRR\_PRR\_TRANSFER. 1.  
 DR PRODOM: PD003042; -; 1.  
 SQ SEQUENCE 406 AA; 45072 MW; 5842CCAD5314DAD0 CRC64;

## Query Match

Best Local Similarity 72.1%; Score 715; DB 13; Length 406;  
 Matches 129; Conservative 26; Mismatches 22; Indels 2; Gaps 1;

QY 1 AOSGCGRGVGRRRYARAKOLVPLLYKQFVGPVPTLGSAGPAEGRVARGSERFDLYPN 60  
 DB 22 ALGCGGR--VIGRRPRPPRLIPLSYKQFLPHVEKTLGASGVEGKIARNSRFELTP 79  
 QY 61 NYNDIIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 120  
 DB 80 NYNDIIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 139  
 QY 121 LAYEGALDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 179  
 DB 140 LAYEGRAVDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 198

## RESULT 3

ID 057404 PRELIMINARY: PRT: 434 AA.

AC 057404  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE SONIC HEDGHO-RELATED PROTEIN.  
 OS Pleurodeles waltlilii (Iberian ribbed newt).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;  
 OC Pleurodeles.  
 NCBI\_TaxID=8319;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Caudit X., Nicolas S., Le Parco Y.  
 RL Submitted (May-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF005332; AAB94412.1; -.  
 DR HSSP: O62226; 1VHH.  
 DR INTERPRO: IPR000320; -.  
 DR INTERPRO: IPR001657; -.  
 DR INTERPRO: IPR001767; -.  
 DR PFAM: PF01079; HIntc. 1.  
 DR PFAM: PF01085; HH\_Signal; 1.  
 DR PRINTS: PR00632; SONICHHOG.  
 DR PRODOM: PD003042; -; 1.  
 SQ SEQUENCE 434 AA; 48421 MW; A9495E367151AE74 CRC64;

## Query Match

Best Local Similarity 74.9%; Score 712; DB 13; Length 434;  
 Matches 128; Conservative 30; Mismatches 16; Indels 2; Gaps 2;

QY 4 CGGGRGVPVGRRRYARAKOLVPLLYKQFVGPVPTLGSAGPAEGRVARGSERFDLYPN 63  
 DB 27 CGGGRG-IGORRRPK-LTPPLAYKQFIPVSEKTLGASGVEGKIARNSRFELTPN 84  
 QY 64 PDIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 123  
 DB 85 PDIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSLHY 144  
 QY 124 EGRALDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 179  
 DB 145 EGRAVDITTSDBDRNRKYGGLARLAVAGFDWYVESRNHVSVKADNSLAVAGG 200

## RESULT 4

ID 09WUP6 PRELIMINARY: PRT: 129 AA.

AC 09WUP6  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE DESERT HEDGHO PROTEIN (FRAGMENT).  
 GN DHH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Garges P.L., Meyer R.A., Jr., Brown C.A., Price D.K.;  
 RT "Desert hedgehog in the rat."  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF148226; AAD31927.1; -.  
 DR HSSP: O62226; 1VHH.  
 DR INTERPRO: IPR000320; -.  
 DR INTERPRO: IPR001657; -.  
 DR PFAM: PF01085; HH\_Signal; 1.  
 DR PRINTS: PR00632; SONICHHOG.  
 FT NON\_TER 1  
 FT NON\_TER 129  
 SQ SEQUENCE 129 AA; 14578 MW; AB33509B3E751319 CRC64;

## Query Match

Best Local Similarity 69.9%; Score 664; DB 11; Length 129;  
 Matches 123; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 OSGGPGRGVGRRRYARAKOLVPLLYKQFVGPVPTLGSAGPAEGRVARGSERFDLYPN 61  
 DB 3 OSGGPGRGVGRRRYARAKOLVPLLYKQFVGPVPTLGSAGPAEGRVARGSERFDLYPN 62  
 QY 62 YNPDIIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 121  
 DB 63 YNPDIIFKDEENSGADRLMTERCKEKERVNALAIAYNNMMPGVRLRYTEGMDDEGHHADSL 122



```
RESULT 8
Q9VCQ4 PRELIMINARY: PRT: 471 AA.
AC Q9VCQ4:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE HH GENE PRODUCT.
GN HH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins K.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktiroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Meyers A.D., Dew I., Dietz S.M.,
RA Dodson K., Doop L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Giodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003742; AAF6102.1; -.
DR HSSP: Q02936; IAT0.
DR FLYBASE: FBgn0004644; hh.
DR INTERPRO: IPR000320; -.
DR INTERPRO: IPR001657; -.
DR INTERPRO: IPR001767; -.
DR PFAM: PF01079; Hlat; 1.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHHOG.
SO SEQUENCE 471 AA; 52150 MW; 8ECD796A92FE7043 CRC64;
```

```
Query Match 63.2%; Score 600.5; DB 5; Length 471;
Best Local Similarity 62.6%; Pred. No. 1.5e-47;
Matches 112; Conservative 28; Mismatches 36; Indels 3; Gaps 2;
```

1 AASCGRGRCGVGRRARRARQVPLVLYKKQFVPCVPTLGCASGAEGRVARGSRFADLYP 60

```
Db 82 AASCGRGRC-LGRHR--ARNLYPLVLYKQTPINLSEYTNASGPLEGIVIRDSKFKDLVP 138
QY 61 NNPDIIFKDEENSGADRLMTTERCKERRVNAIAIVANNMPGVALRVTEGMDGHHADPS 120
Db 139 NNNDLFLFDEBGTGADRLMSKRCRKLAVLAVSWNNKPGIRLVLVTEWDEDDYHHGDS 198
QY 121 LHYEGRALDITTSDRDNKRYGLLARLAVAGFDWYVESRNHVHVSVDKNSLAVRAGC 179
Db 199 LHYEGRAVLIATSDRDSQSKYGLARLAVAGFDWVSYSRRHYCSVKSDSSISHVHG 257

RESULT 9
Q9W6C1 PRELIMINARY: PRT: 138 AA.
AC Q9W6C1:
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
DE SONIC HEDGEHOG PROTEIN (FRAGMENT).
GN SHH.
OS Eleutherodactylus coqui.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;
OC Eleutherodactylus.
OX NCBI_TaxId=57060;
RN [1]
RP SEQUENCE FROM N.A.
RA Carl T.F., Richardson M.K., Olsson L., Schlosser G., Klymkowsky M.W.,
RA Hanken J.;
RT "Differences in vertebrate limb development revealed by studies of the
RT direct developing frog E. coqui.";
RL Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF113403; AAD23436.1; -.
DR HSSP: Q62226; 1VNH.
DR INTERPRO: IPR00320; -.
DR INTERPRO: IPR001657; -.
DR PFAM: PF01085; HH_signal; 1.
DR PRINTS: PRO0632; SONICHHOG.
FT NON_TER 1
FT NON_TER 138
FT TER 138
SO SEQUENCE 138 AA; 15751 MW; FFA156A17E4681F0 CRC64;
```

```
Query Match 63.1%; Score 599; DB 13; Length 138;
Best Local Similarity 75.4%; Pred. No. 4.6e-48;
Matches 104; Conservative 23; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 30 VGVCPERTIGASGPARGARCSERRDLVPNVNPDIIFKDEENSGADRLMTERRKERYN 89
Db 1 IPVVAERTIGASGRGEGKITRNSERKELTPVNSDIIFKDEENTAADRLMTORCKDKLN 60
QY 90 ALAIAVANNMPGVALRVTEGMDGHHADPSLHYEGRALDITTSDRDNKRYGLLARLAVE 149
Db 61 ALAIAVNNMPGVALRVTEGMDGHHSESLHYEGRAVDITTSDRKRYGLLARLAVE 120
QY 150 AGFDWYVESRNHVHVSVDKNSLAVRAGC 167
Db 121 AGFDWYVESKHHHCSV 138

RESULT 10
O61676 PRELIMINARY: PRT: 410 AA.
AC O61676:
DT 01-AUG-1998 (TREMblrel. 07, Created)
DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE HEDGEHOG.
OS Lytechinus variegatus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
```





```
DR PFAM:PF01085; HH_signal; 1.
DR PRINTS: PR00632; SONICHHOG.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 161 AA; 17306 MW; DF060855FD45AEADB CRC64
```

**THIS PAGE BLANK (USPTO)**

---









/organism="Mus musculus"

Kohn, S., Shln, T., Jackson, Y., Cardenas, M., McCann, R.

and Wilson, R.





```

51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI1 67
295 AGGATACCCAAATTTAAAGATCTTGCGTGAACCTTGGACCAACAATGTCGT 344
67 ePheLysAspGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
345 GTTCCGACAGAGAGACGAGCGGCCGACAGATATGACAAAGGAT 394
84 ySlySGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
395 GCCCGGATCGCTCTAAAACCTTGGCCTGATGCGTAATTAATGGCCA 444
101 GLyValArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHISAl 117
445 AAAGTTAA..CTTCTTTTATCCAGCTTAGACAAAACGGAATCCACGG 493
117 aGlnAspSer..LeuHisTyrGluGlyArgValAlaLeuAspIleThrThS 133
494 T...CCGACCCCTGATTCAGAGGTAGAACAGTGCACCTCCCGGAC 540
134 AspArgAspArg..AsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlu 150
541 ACAAGAACCAAGAAAAT...TTCCGAATGCTGGGCTCCGATGCCCTAATG 587
150 laGlyPheAspTrpValTyrTyrGluSerArg 160
588 CCCGGTTGACTGGGTTTTTACCAATCCAA 619

```

seq\_name: gb\_est14:AF201255

seq\_documentation\_block:

LOCUS AF201255 625 bp mRNA EST 29-SEP-2000  
 DEFINITION AF201255 Botryllus schlosseri EST Botryllus schlosseri cDNA, mRNA  
 sequence.  
 ACCESSION AF201255  
 VERSION AF201255.1 GI:10419686  
 KEYWORDS EST.  
 SOURCE Botryllus schlosseri.  
 ORGANISM Botryllus schlosseri.  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;  
 Stolidobranchia; Botryllidae; Botryllus.

REFERENCE 1 (bases 1 to 625)  
 Meyer, R.A., Leonard, P.M., Rosenberg, G.H., Miller, R.D. and Loker

AUTHORS 'E.S.

TITLE Characterization of ESTs from the Colonial Tunicate, Botryllus

JOURNAL schlosseri

COMMENT Unpublished (2000)

CONTACT: Loker ES

Parasitology

Biology, University of New Mexico

Casteretter Hall, Albuquerque, NM 87131, USA

Tel: 505 277 5508

Fax: 505 277 0304

Email: esloker@unm.edu.

FEATURES

1..625 Location/Qualifiers

source /organism="Botryllus schlosseri"

BASE COUNT 174 a 163 c 164 g 124 t

ORIGIN

alignment\_scores:

Quality: 313.50 Length: 162

Ratio: 2.549 Gaps: 5

Percent Similarity: 75.926 Percent Identity: 41.975

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x AF201255 ..

Align seg 1/1 to: AF201255 from: 1 to: 625

```

1 AlaGlnSerGlySGLyProGlyArgGlyProValGlyArgArgTyrAl 17
148 GGGGAGAGCTGC...CGGGGCGGGGCGCGGTTTTCGACCAACACCC 194
17 aArgGlyGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
195 TCACAGACCCAGAGTGGCGGTGATACCGGACGATATTCCAATCTG 244
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
245 CCGAAAACATATCGGGGCTAGCGGGCGCGGACGAGCAATATTCCAAG 294
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI1 67
295 AGGATACCCAAATTTAAAGATCTTGCGTGAACCTTGGACCAACAATGTCGT 344
67 ePheLysAspGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
345 GTTCCGACAGAGAGACGAGCGGCCGACAGATATGACAAAGGAT 394
84 ySlySGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
395 GCCCGGATCGCTCTAAAACCTTGGCCTGATGCGTAATTAATGGCCA 444
101 GLyValArgLeuArgValThrGluGlyTTPAspGluAspGlyHisHISAl 117
445 AAAGTTAA..CTTCTTTTATCCAGCTTAGACAAAACGGAATCCACGG 493
117 aGlnAspSer..LeuHisTyrGluGlyArgValAlaLeuAspIleThrThS 133
494 T...CCGACCCCTGATTCAGAGGTAGAACAGTGCACCTCCCGGAC 537
133 rAspArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGlu 150
538 ACAAGAACCAAGAAAAT...TTCCGAATGCTGGGCTCCGATGCCCTAATG 587
150 laGlyPheAspTrpValTyrTyrGluSerArg 160
588 CCCGGTTGACTGGGTTTTTACCAATCCAA 619

```

seq\_name: gb\_gss31:CNS04F35

seq\_documentation\_block:

LOCUS CNS04F35 926 bp DNA GSS 21-MAY-2000  
 DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone  
 105F04 of library G from Tetraodon nigroviridis, genomic survey  
 sequence.

ACCESSION AL287834.1 GI:8026342

VERSION AL287834.1

KEYWORDS GSS; genome survey sequence.

SOURCE Tetraodon nigroviridis.

ORGANISM Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha;

Holacanthopterygii; Acanthopterygii; Petcomorpha;  
 Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 926)

Roest-Crollius, H., Jallion, O., Dasilva, C., Fizames, C., Fisher, C.,  
 Bouneau, L., Billault, A., Queller, F., Saurin, W., Bernot, A. and

Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the

freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 926)

Roest-Crollius, H., Jallion, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Winkler, P., Brotier, P., Queller, F.,

Saurin, W. and Weissenbach, J.

Human gene number estimate provided by genome wide analysis using

Tetraodon nigroviridis DNA sequence

REFERENCE 3 (bases 1 to 926)

JOURNAL Unpublished

AUTHORS Genoscope.

TITLE Direct Submission  
JOURNAL Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases  
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
source  
1. 926  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="105F04"  
/clone\_1b="c"  
/note="Genoscope sequence ID : COBGI05DC02LPI-end : T7"  
BASE COUNT 214 a 236 c 232 g 242 t 2 others  
ORIGIN

alignment\_scores:  
Quality: 308.00 Length: 77  
Ratio: 4.597 Gaps: 0  
Percent Similarity: 87.013 Percent Identity: 71.429

alignment\_block:  
US-08-900-220C-17\_COPY\_20\_198 x CNS04F35/rev ..

Align seg 1/1 to reverse of: CNS04F35 from: 1 to: 926

```
103 ArgLeuAlValThrGluGlyTyrAspGluAspGlyHisAlaGlnAs 119
   ::::::::::::::::::::|::::::::::::|:::
926 AACCTCGCGGTACCCAGGCGCTGGACACGACACACCTTGGAGCA 877
   ::::::::::::::::::::|::::::::::::|:::
119 pSerLeuHisTyrGluGlyArgAlaLeuAspLleThrSerAspArg 136
   ::::::::::::::::::::|::::::::::::|:::
876 GTCTCTGCATACGAGCGCGCGCTGGACATCACACCTGGACCGGG 827
   ::::::::::::::::::::|::::::::::::|:::
136 spArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaGlnAlaGlyPhe 152
   ::::::::::::::::::::|::::::::::::|:::
826 ACAAGAGCAAAATACGCGACCTGTCCAGCGTGGAGCGCGCTTC 777
   ::::::::::::::::::::|::::::::::::|:::
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
   ::::::::::::::::::::|::::::::::::|:::
776 CACTGGCTCTATATAGTATCCAAAGCTCACCCTCGCTCGTAAAGC 727
   ::::::::::::::::::::|::::::::::::|:::
169 aspAsnSerLeuAlaValArgAlaGlyGly 179
   ::::::::::::::::::::|::::::::::::|:::
726 AGGTACGAGCGCGCGGAGGCGGCGGGGG 696
```

seq\_name: gb\_gss32:CNS05M5G

seq\_documentation\_block:  
LOCUS CNS05M5G 819 bp DNA GSS 26-MAY-2000  
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone  
029G06 of library A from Tetraodon nigroviridis, genomic survey  
sequence.

ACCESSION AL343645  
VERSION AL343645.1 GI:8237415  
KEYWORDS GSS: genome survey sequence.  
SOURCE Tetraodon nigroviridis.  
ORGANISM Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Eurypterygii; Ctenosquamata; Acanthomorpha; Eucanthomorpha;  
Holacanthopterygii; Acanthopterygii; Percomorpha;  
Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE  
AUTHORS Roest-Crolius, H., Jallion, O., Dasilva, C., Fizes, C., Fisher, C.,  
Bonneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and  
Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 819)  
AUTHORS Roest-Crolius, H., Jallion, O., Dasilva, C., Bonneau, L., Fisher, C.,

TITLE Direct Submission  
JOURNAL Submitted (12-Apr-2000) to the EMBL/GenBank/DBJ databases  
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
source  
1. 819  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="029G06"  
/clone\_1b="a"  
/note="Genoscope sequence ID : COA029BD03C1-end : T7"  
BASE COUNT 132 a 239 c 243 g 151 t 54 others  
ORIGIN

alignment\_scores:  
Quality: 222.50 Length: 82  
Ratio: 3.423 Gaps: 2  
Percent Similarity: 79.268 Percent Identity: 62.195

alignment\_block:  
US-08-900-220C-17\_COPY\_20\_198 x CNS05M5G/rev ..

Align seg 1/1 to reverse of: CNS05M5G from: 1 to: 819

```
2 GlnSerCysGlyProGlyArgGlyProValGlyArgArgTyrAlaAr 18
   ::::::::::::::::::::|::::::::::::|:::
402 GAGGCGCTCGGCGCGCGCGCGAG..GCTACGGGAAGCGGGTTCCACGG 354
   ::::::::::::::::::::|::::::::::::|:::
18 GlySerLeuValProLeuTyrLysGlnPheValProGlyValProG 35
   ::::::::::::::::::::|::::::::::::|:::
353 GAGAGTGGCTC...CGCTCGCTTACAGAGTATAGCCCAAGTGGCGG 307
   ::::::::::::::::::::|::::::::::::|:::
35 LuArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGly 51
   ::::::::::::::::::::|::::::::::::|:::
306 AGAAGACTCTGGGGCGCGAGAGATACGAAGGAAATTAACGGCAAN 257
   ::::::::::::::::::::|::::::::::::|:::
52 SerGluArg.PheArgAspLeuValProAspTyrAsnProAspLleLe 68
   ::::::::::::::::::::|::::::::::::|:::
256 TCGAGCGCTTNAAGAGAGCTCACCCCACTACAAACCCGACATCATCT 207
   ::::::::::::::::::::|::::::::::::|:::
68 helYAspGluGluAsnSerGlyAlaAspArgLeuMetThrGlu 82
   ::::::::::::::::::::|::::::::::::|:::
206 TCAAGGATGAGGAGAACACAGGTGCGACCGCTGATGCGGAG 163
```

seq\_name: gb\_est48:AW603741

seq\_documentation\_block:  
LOCUS AW603741 291 bp mRNA EST 23-MAR-2000  
DEFINITION CM2-CN0044-190100-073-f10 CN0044 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW603741  
VERSION AW603741.1 GI:7308482  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS HCGP <http://www.ludwig.org.br/OESTES>.  
TITLE The FAPESP/LICR Human Cancer Genome Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/getnum2.pl?cl=CM2&cl2=CM2-CN0044-190100-073-f10&cl3=2000-01-19&cl4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 173.  
 High quality sequence stop: 173.

## FEATURES

## source

1..291  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1ib="CN0044"  
 /dev\_stage="Adult"  
 /note="Organ: colon.normal; Vector: puc18; site\_1: Sma1; site\_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 45 a 93 c 96 g 57 t  
 ORIGIN

## alignment\_scores:

Quality: 195.00 Length: 60  
 Ratio: 3.824 Gaps: 0  
 Percent Similarity: 85.000 Percent Identity: 61.667.

## alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x AM603741 ..

Align seg 1/1 to: AM603741 from: 1 to: 291

```

120 SerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgAs 136
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
12 TCCTCGCTTATGACGCGCGCGGTGACATCACACATCACAGCCGCGA 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
136 PARGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPheA 153
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
62 CCGCATACGATGACTGCTGGCGCGCTTGCGACCTGAGCGCGCTTTTG 111
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
153 sPTpValTyrTyrGluSerArgAsnHisValHisValSerValLysAla 169
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
112 ACTGGTGTCTTACGACAAAGGCCCGCATGTTCTCCGTCTGCTCC 161
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
170 AspAsnSerLeuAlaValArgAlaGlyGly 179
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
162 GAGCACATCGCGCGCGACCAAGACGGCGGC 191
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: gb\_est15:AI015155

## seq\_documentation\_block:

LOCUS AI015155 455 bp mRNA EST 27-AUG-1998  
 DEFINITION O70h05.s1 Soares\_total\_fetus.Nb2HF8\_9w Homo sapiens cDNA clone  
 IMAGE:1622169 3' similar to TR:Q98938 Q98938 INDIAN HDGEBHOG  
 PROTEIN:,, mRNA sequence.

ACCESSION AI015155  
 VERSION AI015155.1 GI:3229491

## KEYWORDS

## SOURCE

## ORGANISM

human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 455)  
 NC1-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert Length: 615 Std Error: 0.00  
 Seq primer: -40m3 fwd. Et from Amersham  
 High quality sequence stop: 392.

## FEATURES

## source

1..459  
 Location/Qualifiers

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1ib="IMAGE:1622169"  
 /clone\_1ib="Soares\_total\_fetus.Nb2HF8\_9w"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
 was prepared from mRNA obtained from pooled 8-9 week  
 (total) fetus material with a Not I - oligo(dT) primer (5'  
 TGTTCATCTGACATGAGTCGAGCGCGCCCTTAAATTTTCTTTTCTTTT 3').  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia). digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 106 a 127 c 133 g 93 t  
 ORIGIN

## alignment\_scores:

Quality: 180.00 Length: 61  
 Ratio: 3.673 Gaps: 3  
 Percent Similarity: 80.328 Percent Identity: 67.213

## alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x AI015155/rev ..

Align seg 1/1 to reverse of: AI015155 from: 1 to: 459

```

5 GTPProGlyArgGlyProValGlyArgArg.ArgTyrAlaArgLysInt 21
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
459 GGGCGGGGTGGG...GTGGTGGGACCGCGCGCGACCGCCAGCA...C 416
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 euValProLeuLeuTyrLysGlnPheValProGlyValProGlyArgThr 37
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
415 TCGTGGCGGTGCGCTTACACAGCAGTTTCAGCCCAATGTGCCGAGAGACC 366
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
38 LeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgLysSerGluAr 54
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
365 CTGGCGCCCGACGCGACGCTTGAAGGCAAGATCGCTCGACGTCCGAGCG 316
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
54 gPheArgAspLeuValProAsnTyrAsnPro 64
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 CTTCATGAGCTCACCCGCCCAATTACATCCA 285
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

seq\_name: gb\_est38:AV604922

## seq\_documentation\_block:

LOCUS AV604922 555 bp mRNA EST 30-AUG-2000  
 DEFINITION AV604922 Bos taurus kidney fetus Bos taurus cDNA clone EIKI025F10  
 5', mRNA sequence.

ACCESSION AV604922  
 VERSION AV604922.1 GI:9735295

## KEYWORDS

## SOURCE

## ORGANISM

cow.  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 555)  
 Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and

TITLE  
JOURNAL  
COMMENT

Suzuki, H.  
bovine CDNA sequencing  
Unpublished (2000)  
Contact: Yoshikazu Sugimoto  
Animal Genetics Division  
Shirakawa Institute of Animal Genetics  
Odaoka, Nishigo, Nishi-Shirakawa, Fukushima 961-8061, Japan  
Tel: 81-248-25-5641  
Fax: 81-248-25-5725  
Email: kazusugi@cocoa.ocn.ne.jp  
Single pass sequencing.  
This clone was obtained from a polyA-deleted cDNA library.

## FEATURES

## source

Location/Qualifiers  
1..355  
/organism="Bos taurus"  
/db\_xref="taxon:9913"  
/clone="E1K1025F10"  
/clone\_lib="Bos taurus kidney fetus"  
/tissue\_type="kidney"  
/dev\_stage="fetus"  
/lab\_host="DH10B"  
/note="Vector: pZ1; Site\_1: SalI; Site\_2: NotI; Poly A  
was deleted from a NotI site"  
BASE COUNT 98 a 184 c 175 g 98 t  
ORIGIN

## alignment\_scores:

Quality: 165.00 Length: 46  
Ratio: 4.024 Gaps: 1  
Percent Similarity: 89.130 Percent Identity: 73.913

## alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x AV604922 ..

Align seg 1/1 to: AV604922 from: 1 to: 555

135 ArgAspArgAsnLysTyGlyLeuLeuAla..ArgLeuAlaValGluAla 150  
|||||  
10 CGCGACCGCATTAAGTACGACGCTGACGACCGCTTGCGAGTGGAGCGCC 59  
151 GlyPheAspTPrpValTyTyGluSerArgSnhSValHISValSerVa 167  
|||||  
60 GCGTTCAGCTGGGTGTATTACGAGTCCAGGCCAGCGTTCCTCGT 109  
|||||  
167 LlysAlaAspAsnSerLeuAlaValArgAlaGlyGly 179  
|||||  
110 CAAGTCCGAGCAGCTAGCGCCGACGACGAGAGTGGC 146

seq\_name: gb\_est54:BB030536

## seq\_documentation\_block:

LOCUS BB030536 239 bp mRNA EST 23-JUN-2000  
DEFINITION BB030536 RIKEN full-length enriched, adult male thymus Mus musculus  
cDNA clone 5830455115 3' similar to AF148226 Rattus norvegicus  
desert hedgehog protein (dhn) mRNA, mRNA sequence.

ACCESSION BB030536  
VERSION BB030536.1 GI:8385287

## KEYWORDS

## SOURCE

## ORGANISM

house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 239)

## REFERENCE

## AUTHORS

Kono, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci,  
P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,  
Hirozane, T., Horii, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,  
Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,  
Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurihara, C., Kusakabe, M.,  
Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,  
Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata,  
Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,  
Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomioka, N., Toya

TITLE  
JOURNAL  
COMMENT

T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,  
Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino,  
M., Muramatsu, M. and Hayashizaki, Y.  
RIKEN Mouse ESTs (Kono, H., et al.)  
Unpublished (2000)  
Contact: Yoshinori Hayashizaki  
Genome Exploration Research Group, Life Science Tsukuba Center,  
Genome Science Laboratory  
The Institute of Physical and Chemical Research (RIKEN), Genomic  
Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan  
Tel: +81-298-36-9013  
Fax: +81-298-36-9038  
Email: genome-res@rtc.riken.go.jp,  
URL: http://genome.rtc.riken.go.jp/  
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki,  
N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Thermolabile and thermoactivation of thermolabile enzymes by  
trehalose and its application for the synthesis of full length  
cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,  
Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki,  
Y. and Hayashizaki, Y.  
Automated filtration-based high-throughput plasmid preparation  
system. Genome Res. 9 (5), 463-470 (1999)  
Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
19-44 (1999)  
Please visit our web site (<http://genome.rtc.riken.go.jp>) for  
further details.

## FEATURES

## source

Location/Qualifiers  
1..239

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="5830455115"  
/clone\_lib="RIKEN full-length enriched, adult male thymus"  
/sex="male"  
/tissue\_type="thymus"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3'], cDNA was  
transcribed by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. cDNA went through one round of normalization  
to Rot = 10.0 and subtraction to Rot = 100.0. Second  
strand cDNA was prepared with the primer adapter of  
sequence [5' GAGAGAGAGATTCGAGTAAATTAATTAATCCCCCCCCCC  
3']. cDNA was cloned into the XhoI and BamHI sites.  
Vector: a modified Bluescript KS(+) after bulk excision  
from lambdaBda FLC I. Cloning sites, 5' end: SalI; 3' end:  
BamHI."

BASE COUNT 60 a 72 c 48 g 59 t  
ORIGIN

## alignment\_scores:

Quality: 160.00 Length: 39  
Ratio: 4.444 Gaps: 0  
Percent Similarity: 92.308 Percent Identity: 71.795

## alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x BB030536/rev ..

Align seg 1/1 to reverse of: BB030536 from: 1 to: 239

```

83 ArgCysIysGluArgValAsnAlaLeuAlaIleAlaValMetAsnMetTr 99
|||||
138 CGTTGCAAGAGAGGGGTGAATGCTTATCCATCGCGGTGATTAATATATGTG 89
|||||
99 pProGlyValArgLeuArgValThrGluGlyTTPaspGluAspGlyHisH 116
|||||
88 GGGCGGAGTACGGGTAGTGAATGAAGGGGTGGGAAAAAGATGGGCACC 39
|||||
116 tSAlaGlnAspSerLeu 121
|||||
38 ATGCTTAGGAACCTCTC 22

```



CC lesions of carcinomas. The methods can also be used to counteract  
CC the effects of ageing on skin.

SQ Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

# alignment\_scores:

Quality: 950.00 Length: 179  
Ratio: 5.307 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x X25624 ..

Align seg 1/1 to: X25624 from: 1 to: 1190

```

1  ALaGInSerCgLyPrOgLyArGgLyProVaLgIaRgArGArGTYrAl 17
   |||||||
58  GCGCAGAGCTGCGGGCGCGGGCGGCGGTGGCGCGCGCGCTATGC 107
   |||||||
17  aArGlySGInLeuValProLeuLeuTYrLySGInPheValProGlyValP 34
   |||||||
108  GCGCAGAGAGCTGCGCGCTACTCTACAAACAAATTGTGCGCGCGCTGC 157
   |||||||
34  roGluArGThLeuGlyAlaSerGlyProAlaGluGlyArGValAlaArg 50
   |||||||
158  CAGAGCGGACCTGGGCGCCAGTGGCGCGAGGAGGAGGAGGTGGCAAGG 207
   |||||||
51  gLySerGluArGpheaRgAspLeuValProAsnTYrAsnProAspIleI 67
   |||||||
208  GGCCTCCGAGCGCTTCCGGGACCTCGTGCCCAACTACAAACCCGACATCAT 257
   |||||||
67  ePhaLyAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
   |||||||
258  CTTCAAGAGTAgAGAGAACACTGAGCGCGAGCGCTGTATGACGAGCGCTT 307
   |||||||
84  ySLySGluArGValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
   |||||||
308  GCAAGGAGAGGCTGACGCTTTGGCCATTGGCGCTGATGATGACATGTGGCCC 357
   |||||||
101  gLyValArGLeuArGValThrGluGlyTrpAspGluAspGlyHisHisAl 117
   |||||||
358  GGAAGTGGCGCTACGAGTCACTGAGGCGGTGGAGACGAGCGCCACACAGC 407
   |||||||
117  aGInAspSerLeuHisTYrGluGlyArgAlaLeuAspIleThrThrSerA 134
   |||||||
408  TCAGGATTCACTCCACTACGAAAGCGCGTCTTGACATCACTACGCTGTG 457
   |||||||
134  sPaRgAspARgAsnLySTyrgLyLeuLeuAlaArGLeuAlaValGluAla 150
   |||||||
458  ACCCGGAGCGGACAAAGATAGTGGTTGCTGGCGCGCTGCGAGTGGAGGCC 507
   |||||||
151  gLyPheaSPTrpValTYrTYrGluSerARgAsnHisValHisValSerVa 167
   |||||||
508  GCGTTTCGACTGGGTCTACTACGAGTCCCGCAACACAGTCCACGTCGTGGGT 557
   |||||||
167  lLyAlaAspAsnSerLeuAlaValArGAlaGlyGly 179
   |||||||
558  CAAAGCTGATAACTCACTGGCGGCTCCGGCGCGCGCGG 594
   |||||||
seq_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.X25105
seq_documentation_block:
ID X25105 standard; cDNA: 1190 BP.
XX
AC X25105;
XX
DT 05-JUL-1999 (first entry)
XX
DE Human Desert hedgehog protein Dhh cDNA.
XX
KW Desert'hedgehog'; Dhh gene; human; hedgehog therapeutic;
KW ptc therapeutic; patched; signal transduction; muscle atrophy;

```

KW cachexia; muscular myopathy; myoblastic sarcoma; therapy; ds.

XX Homo sapiens.

XX MO9910004-A2.

XX 04-MAR-1999.

XX 28-AUG-1998; 98MO-US17922.

XX 29-AUG-1997; 97US-0057394.

XX (ONTO-) ONTOGENY INC.

XX Bladgen CS, Currie PD, Hughes SM, Ingham PW;

XX WPI; 1999-243557/20.

XX P-PsDB; Y05517.

XX A new method to regulate muscle growth

XX Disclosure; Page 106-108; 130pp; English.

CC This nucleotide sequence comprises a coding region for the human  
CC desert hedgehog protein Dhh (see Y05517). The invention relates to  
CC a method for modulating the formation and/or maintenance of muscle  
CC tissue by ecotopically contacting muscle cells, especially muscle  
CC stem/progenitor cells, in vitro or in vivo, with a hedgehog  
CC therapeutic (i.e. hedgehog polypeptides and gene therapy  
CC constructs) or ptc therapeutic (i.e. a small organic molecule that  
CC mimics the effect of hedgehog proteins on patched signalling, or  
CC activates or potentiates patched signalling) in an amount effective  
CC to alter the growth state of the treated cells. Also claimed is a  
CC method for treatment or prevention of disorders of, or surgical or  
CC cosmetic repair of, such muscle tissues, by administering a  
CC hedgehog polypeptide or ptc therapeutic. The disorder may be  
CC muscle atrophy, in particular skeletal muscle atrophy or cardiac  
CC muscle atrophy, cachexia, or muscular myopathy (all claimed). The  
CC hedgehog polypeptide or ptc therapeutic can inhibit growth of  
CC myoblastic-derived tissue to provide treatment of hyperplastic or  
CC neoplastic growth of muscle tissue such as in myoblastic sarcoma  
CC (also claimed). The hedgehog therapeutic preferably comprises at  
CC least a bioactive extracellular portion of a hedgehog protein (see  
CC Y05510-19) encoded by a vertebrate hedgehog gene (see X25098-107),  
CC especially a human hedgehog gene.

SQ Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

# alignment\_scores:

Quality: 950.00 Length: 179  
Ratio: 5.307 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

# alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x X25105 ..

Align seg 1/1 to: X25105 from: 1 to: 1190

```

1  ALaGInSerCgLyPrOgLyArGgLyProVaLgIaRgArGArGTYrAl 17
   |||||||
58  GCGCAGAGCTGCGGGCGCGGGCGGCGGTGGCGCGCGCGCTATGC 107
   |||||||
17  aArGlySGInLeuValProLeuLeuTYrLySGInPheValProGlyValP 34
   |||||||
108  GCGCAGAGAGCTGCGCGCTACTCTACAAACAAATTGTGCGCGCGCTGC 157
   |||||||
34  roGluArGThLeuGlyAlaSerGlyProAlaGluGlyArGValAlaArg 50
   |||||||
158  CAGAGCGGACCTGGGCGCCAGTGGCGCGAGGAGGAGGTGGCAAGG 207
   |||||||
51  gLySerGluArGpheaRgAspLeuValProAsnTYrAsnProAspIleI 67
   |||||||

```







DE Human Desert hedgehog Dhh cDNA.  
 XX  
 KW Desert hedgehog; Dhh; human; excitotoxicity; Parkinson's disease;  
 KM Huntington's disease; neuronal degeneration; neuroprotective;  
 KM dopaminergic; GABAergic; substantia nigra; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200035948-A1.  
 PD 22-JUN-2000.  
 PF 03-DEC-1999; 99WO-US28721.  
 PR 03-DEC-1998; 98WO-US25676.  
 PR 27-JAN-1999; 99US-0238243.  
 PR 03-JUN-1999; 99US-0325602.  
 PA (BIOJ) BIOGEN INC.  
 PI (ONTO-) ONTOGENY INC.  
 XX Galdes A, Mahanthappa N;  
 PI WPI: 2000-431570/37.  
 DR P-PSDB; Y95288.  
 XX  
 PT Treating disorders involving exotoxicity, e.g. trauma, hypoglycemia,  
 PT senile dementia and Korsakoff's disease, by using lipophilic modified  
 PT hedgehog polypeptide -  
 XX  
 PS Disclosure; Page 145-147; 174pp; English.  
 XX  
 CC The present sequence of that of cDNA coding for human Desert  
 CC hedgehog (Dhh) protein (see Y95288). The invention relates to a  
 CC method for promoting the survival and/or functional performance  
 CC of neuronal cells, especially substantia nigra, dopaminergic or  
 CC GABAergic neurons that are susceptible to exotoxicity, by  
 CC contacting the cells, in vitro or in vivo, with a lipophilic  
 CC (e.g. cholesterol) modified hedgehog polypeptide. The method is  
 CC used to treat or prevent Parkinson's disease, Huntington's disease,  
 CC domoic acid poisoning, spinal cord trauma, hypoglycemia, mechanical  
 CC trauma to the nervous system, senile dementia, Korsakoff's disease,  
 CC schizophrenia, AIDS dementia, multi-infarct dementia, mood  
 CC disorders, depression, chemical toxicity, neuronal damage  
 CC associated with uncontrolled seizures such as epileptic seizures,  
 CC neuronal injury associated with HIV and AIDS, neurodegeneration  
 CC associated with Down's syndrome, neuropathic pain syndrome,  
 CC olivopontocerebral atrophy, amyotrophic lateral sclerosis,  
 CC mitochondrial abnormalities, Alzheimer's disease, hepatic  
 CC encephalopathy, Tourette's syndrome and drug addiction (all  
 CC claimed). The lipophilic modified hedgehog polypeptide is also  
 CC useful for promoting survival and/or functional performance of  
 CC neuronal cells susceptible to exotoxicity. Hedgehog gene  
 CC constructs can be used for recombinant production of hedgehog  
 CC polypeptides, to design probes and primers, and in gene therapy  
 CC protocols to deliver nucleic acids encoding a neuroprotective  
 CC form of hedgehog polypeptide.  
 XX  
 SO Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

Alignment scores:  
 Quality: 950.00 Length: 179  
 Ratio: 5.307 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

Alignment block:  
 US-08-900-220C-17\_COPY\_20\_198 x A27883  
 Align seg 1/1 to: A27883 from: 1 to: 1190

1 AAGinsrCysGlyProGlyArgGlyProValGlyArgArgArgTyrAl 17  
 ||||||||||||||||||||||||||||||||||||||||||||

58 GCCCAGAGCTGGCGGGCCGGGGCCGGGCTTGGCGGGCCGGCTATATC 107  
 17 aarglysginleuValProleuleuTyrLysGlnPheValProGlyValp 34  
 ||||||||||||||||||||||||||||||||||||||||||||  
 108 GCGCAAGAGAGCTGTCGCCCTACTACTACAAGCAATTTGGCCCGGGTCC 157  
 34 rogluarGThrleuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50  
 ||||||||||||||||||||||||||||||||||||||||||||  
 158 CAGAGCGAGACCTGGCGGCGGAGTGGGCGCAGCGAGGGAGGGTGGCAAG 207  
 51 GlyserGluArgPheArgAspleuValProAsnTyrAspAspIleI 67  
 ||||||||||||||||||||||||||||||||||||||||||||  
 208 GGCTCCGAGCGCTCCGGGACCTCGTCCCACTACAAACCCGACCTTCAT 257  
 67 ephelysAspGluGluAsnSerGlyAlaAspArgleuMetThrGluArg 84  
 ||||||||||||||||||||||||||||||||||||||||||||  
 258 CTTCAAGAGATGAAGAGAACAGTGAGCCGACCCCTGATGACCGAGCTT 307  
 84 yslYsgLuarGValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100  
 ||||||||||||||||||||||||||||||||||||||||||||  
 308 GCAAGGAGAGGGTGAACGCTTGGCCATTCGCGTGAACAATGTGGCCC 357  
 101 GlyValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAl 117  
 ||||||||||||||||||||||||||||||||||||||||||||  
 358 GGAGTGGCTGACTGAGTACTGAGGGCTGGAGAGGAGCGCCACCCAGCC 407  
 117 aglnAspSerleuHisTyrGluGlyArgAlaLeuAspIleThrSera 134  
 ||||||||||||||||||||||||||||||||||||||||||||  
 408 TCAGGATTCACCTCCACTACGAGGCCGCTTGGACATCAGTACTGCG 457  
 134 sPArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150  
 ||||||||||||||||||||||||||||||||||||||||||||  
 458 ACCGCGACCCGCAACAGTATGGCTTCGCGCGCCGACAGTGAAGCC 507  
 151 GlyPheAspTyrValTyrTyrGluSerArgAsnHisValHisValSera 167  
 ||||||||||||||||||||||||||||||||||||||||||||  
 508 GGCTTGACTGGGTACTGACGAGTCCCGCAACACAGTCCACGTGCGGT 557  
 167 llysaAlaAspAsnSerleuAlaValArgAlaGlyGly 179  
 ||||||||||||||||||||||||||||||||||||||||||||  
 558 CAAGCTGATTACTCACTGGCGGTCCGGCGCGCGCGC 594

seq.name: /SID52/gcdata/geneseq/geneseqn/NA2000.DAT:A30281

seq.documentatn\_block:  
 ID A30281 standard; cDNA; 1190 BP.  
 XX  
 AC A30281;  
 XX  
 DT 11-SEP-2000 (first entry)  
 XX  
 DE Partial human Dhh coding sequence.

Human; desert hedgehog; Dhh; neuromuscular disorder; neuropathy;  
 Guillian-Barre syndrome; peripheral neuropathy; diabetes; alcoholism;  
 Chronic inflammatory demyelinating polyneuropathy; CIPD;  
 gene therapy; infection; inflammation; hereditary neuropathy;  
 Charcot-Marie-Tooth disease; vasculitis; lung cancer; tumour;  
 multiple myeloma; nutritional imbalance; kidney disease;  
 hypothyroid neuropathy; trauma; Refsum's disease; Abetalipoproteinemia;  
 Tangier disease; Krabbe's disease; Metachromatic leukodystrophy;  
 Fabry's disease; CMT; GBS; Dejerine-Sottas syndrome; acute neuropathy;  
 Amyotrophic lateral sclerosis; ALS; Miller-Fisher syndrome; amyloidosis;  
 Hereditary sensory neuropathy Type II; HSN II; B-cell lymphoma;  
 Waidenström's Macroglobulinaemia; Chronic Lymphocytic Leukemia;  
 neuroprotective; cytoprotective; patched-mediated signal transduction;  
 ss.  
 KW  
 KM  
 XX  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH 1..1188  
 FT /\*tag= a



DR WPI: 2000-271252/23.  
P-PSDB: Y70683.

XX Modulation of lung tissue or cell growth rate used for treating or  
PT preventing damage to lung tissue comprises ectopically contacting  
PT tissue with hedgehog therapeutic, patched therapeutic or fibroblast  
PT growth factor-10 -

Claim 15; Page 106-109; 143pp; English.

The patent discloses a method for modulating the growth state of  
CC epithelial or mesenchymal cells of the lung, by ectopically contacting  
CC the tissue with a therapeutic agent, that can effectively alter the rate  
CC of proliferation of cells. This agent can be selected from hedgehog (hh),  
CC patched (ptc) or fibroblast growth factor (fgf)-10 therapeutics. It  
CC involves a direct or indirect antagonism of patched-mediated regulation  
CC of gene expression. This method is useful for the treatment or prevention  
CC of lung diseases, like cancer, cystic fibrosis, bronchopneumocystosis,  
CC bronchitis, bronchospasm, sarcoidosis, silicosis, eosinophilic granuloma,  
CC ankylosing spondylitis, emphysema, tuberculosis, respiratory distress  
CC syndrome, allergic rhinitis, asthma, pulmonary fibrosis and primary  
CC pulmonary hypertension. It is also used to control wound healing or other  
CC reformation processes in the lung and augment lung transplantation. The  
CC present DNA sequence is the human desert hedgehog (Dhh) gene, essential  
CC for development of the respiratory system. Hedgehog polypeptides can be  
CC used to control the formation and/or maintenance of the lung tissue.

Sequence 1190 BP; 176 A; 375 C; 424 G; 215 T; 0 other;

alignment\_scores:  
Quality: 950.00 Length: 179  
Ratio: 5.307 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x Z52264 ..

Align seg 1/1 to: Z52264 from: 1 to: 1190

```

1 AAGInserCyGlyProGlyArgGlyProValIGlyArgArgGlyTyrAl 17
  |||||||
58 GCCCAGAGCTGGGGCCGGGGCCGGGGCGGTGGCGGGCGCGCTATCC 107
  |||||||
17 aargtysglnleuvalproleuleutyrtylsglnphevalproglyvalp 34
  |||||||
108 GCGCAGAGCAGCTGCGCCCTACTCTACAGCAATTTGCGCCGGCGTCC 157
  |||||||
34 rogluarthrleuglalaaserglyproalaglulglaryalalarg 50
  |||||||
158 CAGACCGGACCTGGGGCCGACGTCGAGCGAGGCGGCGTGGCCAAAG 207
  |||||||
51 glysergluarpheargaspheulvalproasnytyrasproasplell 67
  |||||||
208 GGCTCCGAGCGCTTCGGGACCTCGGCCCACTCAACCCGACCTCAT 257
  |||||||
67 ephelysaspGluGluAsnserGlyAlaAspArgLeuMetThrGluArgC 84
  |||||||
258 CTTCAAGATGAGGAGAACAGTGGAGCCGACCGCTGATGACGAGCGTT 307
  |||||||
84 yslsgluarValasnaIalaValaIleValaMetAsnMetTrrPro 100
  |||||||
308 GCAAGAGAGCGGTGAACGCTTTGGCCATTCGCGTATGAAATGTCGCC 357
  |||||||
101 GylValArgLeuArgValThrGluGlyTrrPaspGluAspGlyHisHisAl 117
  |||||||
358 GGAAGTCCGCTTACGAGTACTAGAGGCTGGAGCGAGAGCGCCACCACCC 407
  |||||||
117 aGlnAspSerLeuHisTyrGluGlyAlaArgAlaLeuAspIleThrTrSerA 134
  |||||||
408 TCAGGATTCACCTCACTACAGAGCGCGTCTTTGGACATCCTACAGCTCG 457
  |||||||
134 sPArgAspArgAsnIlyTyrGlyLeuValAlaArgLeuAlaValGluAla 150
  |||||||

```

```

|||||
458 ACCGGCAGCCGCAAGATGAGTGGTGTCTGGCGCGCTCGCACTGGAAGCC 507
  |||||||
151 GlyPheAspTrpValTyrTyrGluSerArgAsnHisValHisSerVa 167
  |||||||
508 GGCTTCGAGTGGGTCTACTACAGAGTCCCGCAACGACGTCCAGTGTGGT 557
  |||||||
167 llysaIaAspAsnSerLeuValaValaArgAlaGlycyl 179
  |||||||
558 CAAGCTGTATTAACCTCACTGCGCGTCCGGCGGGCGGC 594
  |||||||

```

seq\_name: /STS2/gcdata/geneseq/geneseqn/MA1999.DAT:200317

seq\_documentation\_block:

ID 200317 standard; DNA: 1820 BP.

XX 200317;

DT 22-OCT-1999 (first entry)

XX Human desert hedgehog (Dhh-H) protein encoding DNA.

DE Desert hedgehog; Dhh-H; male infertility; menopause; human; CNS;

KW spinal growth injury; spinal cord regeneration; Parkinson's disease;

KW tumour growth; central nervous system disorder; Alzheimer's disease; ss.

XX Homo sapiens.

FH Key Location/Qualifiers

FT CDS 180..1370

FT /tag= a /transl\_except= (pos:438..440, aa:Xaa)

FT /product= "Dhh-H protein"

FT /note= "Xaa = unknown"

FT sig\_peptide 180..245

FT /\*tag= b

FT /note= "putative signal sequence"

FT mat\_peptide 246..1367

FT /\*tag= c

XX W09939725-A1.

XX 12-AUG-1999.

XX 04-FEB-1999; 99WO-US02440.

XX 06-JUL-1998; 98US-0091843.

XX 06-FEB-1998; 98US-0073878.

XX 02-MAR-1998; 98US-0076553.

XX (ELIL ) LILLY & CO ELI.

XX Edmonds BT, King MW, Smith RC;

XX WPI: 1999-508454/42.

XX P-PSDB: Y27264.

XX New human desert hedgehog (Dhh-H) protein useful for, e.g. treating

XX male infertility and menopause

XX Example 1; Page 58-61; 69pp; English.

This DNA encodes a human desert hedgehog (Dhh-H) protein. The Dhh-H  
CC protein can be expressed by standard recombinant methodology.  
CC Compositions containing Dhh-H form methods of treating male infertility  
CC and menopause in humans. The compositions are additionally useful for  
CC treating spinal growth injury, promoting spinal cord regeneration,  
CC inhibiting tumour growth and treating central nervous system (CNS)  
CC disorders in humans. Such CNS disorders includes Alzheimer's disease and  
CC Parkinson's disease. Dhh-H is also useful for identifying compounds that  
CC bind to the protein. The use of Dhh-H for producing antibodies that are  
CC useful in diagnosis and therapy.

alignment\_scores: Quality: 950.00 Length: 179  
 Ratio: 5.307 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-08-900-220C-17\_COPY\_20\_198 x 200317 ..  
 Align seg 1/1 to: 200317 from: 1 to: 1820

```

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAla 17
|||||
237 GCCCAGAGCTGGCGGGCGGGGGCGGTTGGCCGGCGCGCTATGC 286
17 aArgLysGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
|||||
287 GCCGACGACACTGCTGGCGCTACTCTACAAACATTTGTGCGCGCGG 336
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
337 CAGAGCGAGACCTGGGGCCAGTGGCCAGCGAGGGAGGGGTGCAAG 386
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
387 GCCTCCGAGCGCTTCGGGACCTGCTGCCAACATCAACCCGACATCAT 436
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
437 CTTYAAGATGATGAGAACACAGTGGAGCCGACCTTATGACCAACGCTT 486
84 yslYsgLuarGValAsnAlaLeuAlaIleAlaValMetAsnMetTyrPro 100
|||||
487 GTAAGAGCGGGTGAACGCTTGGCCATGTGCCGTATGAACATGTGGGCC 536
101 GlyValArgLeuArgValThrGluGlyTyrAspGluAspGlyHisHisAla 117
|||||
537 GGAGTCGCGCTACGAGTGAAGCTGGAGCGGAGCAGACGCCACCAAGC 586
117 aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerA 134
|||||
587 TCGAGATTCACTCCACTACGACGAGCGCTGTTGGACATCATCATGCTG 636
134 sPaRgsAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
|||||
637 ACCGCGACCGCACCAAGTATGGTGTGCTGCGCGCTCGCAGTGAAGCC 686
151 GlyPheAspTyrValTyrTyrGluSerArgAsnHisValHisValSerVa 167
|||||
687 GCGCTTGACACTGGCTACTACGAGTCCCGCAACCACTCCACGTCGTGT 736
167 llyValAspAsnSerLeuAlaValArgAlaGly 179
|||||
737 CAAAGCTGATTACTACTGCGGTCCGGGGGGGGG 773
737 CAAAGCTGATTACTACTGCGGTCCGGGGGGGGG
  
```

seq\_name: /SID2/gcdata/geneseq/geneseqn/NA2000.DAT:A58931

seq\_documentation\_block:  
 ID A58931 standard; DNA; 1820 BP.  
 XX A58931;  
 AC  
 XX  
 DT 20-OCT-2000 (first entry)  
 DE DNA encoding a desert hedgehog protein.  
 XX  
 DE Human; desert hedgehog protein; Dh-H; male menopause; male virility;  
 XX testosterone level; male infertility; erectile dysfunction;  
 KW tumour growth; spinal regeneration; ss.  
 XX  
 XX Homo sapiens.

---

XX Key Location/Qualifiers  
 FH CDS . 180..1370  
 FT  
 FT  
 FT  
 XX /tag= a  
 PN /product= "desert hedgehog protein"  
 XX  
 XX WO200041537-A2.  
 PD  
 PD 20-JUL-2000.  
 XX  
 XX 12-JAN-2000; 2000MO-US00706.  
 XX  
 XX 14-JAN-1999; 99US-0115931.  
 PR  
 XX (ELIL ) LILLY & CO ELI.  
 XX  
 XX Edmonds BT;  
 XX  
 XX WPI: 2000-475908/41.  
 DR P-PSDB; B07528.  
 DR  
 XX  
 XX Method of treating male menopause, increasing testosterone levels and  
 PT enhancing male virility uses administration of a human homolog of  
 PT Desert hedgehog protein -  
 XX  
 XX Example 1; Page 45-49; 52pp; English.  
 PS  
 XX The present sequence encodes a human desert hedgehog protein (Dh-H).  
 CC The polypeptide is used for treating male menopause. Dh-H is used  
 CC for increasing testosterone levels and enhancing male virility.  
 CC Dh-H can also be used to produce antibodies which are then useful for  
 CC diagnosis to screen for Dh-H or potential modulators of Dh-H e.g. In  
 CC competitive displacement assay or in therapeutic applications.  
 CC Dh-H protein is used to treat male infertility, male menopause,  
 CC erectile dysfunction, diminished virility, tumour growth or spinal  
 CC regeneration.  
 XX  
 XX Sequence 1820 BP; 336 A; 520 C; 620 G; 343 T; 1 other:

alignment\_scores: Quality: 950.00 Length: 179  
 Ratio: 5.307 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
 US-08-900-220C-17\_COPY\_20\_198 x A58931 ..  
 Align seg 1/1 to: A58931 from: 1 to: 1820

```

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAla 17
|||||
237 GCCCAGAGCTGGCGGGCGGGGGCGGTTGGCCGGCGCGCTATGC 286
17 aArgLysGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValP 34
|||||
287 GCCGACGAGCTGCTGGCGCTACTCTACAAACATTTGTGCGCGCGGCTGC 336
287 GCGCAAGCAGCTGCTGGCGCTACTCTACAAACATTTGTGCGCGCGGCTGC 336
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
|||||
337 CAGAGCGAGACCTGGGGCCACTGCGCGAGCGAGGGGGGTGGCCAAAG 386
51 GlySerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
|||||
387 GCCTCCGAGCGCTTCGGGACCTGCTGCCAACATCAACCCGACATCAT 436
67 ePheLysAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
|||||
437 CTTYAAGATGATGAGAACACAGTGGAGCCGACCTTATGACCAACGCTT 486
84 yslYsgLuarGValAsnAlaLeuAlaIleAlaValMetAsnMetTyrPro 100
|||||
487 GTAAGAGCGGGTGAACGCTTGGCCATGTGCCGTATGAACATGTGGGCC 536
  
```







XX This cDNA sequence codes for a precursor (see W79593) of a novel  
 CC human Desert hedgehog protein (see also W79593). Human Desert  
 CC hedgehog protein DNA (HuHh) was originally identified following  
 CC screenings of human cell lines by PCR using primers (see V62402-03)  
 CC based on the mouse Desert hedgehog gene. These screenings  
 CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is  
 CC derived from the plasma cell of a leukemia patient, expressed a  
 CC specific gene at an elevated level. Sequencing confirmed it to  
 CC be a novel human gene showing homology to the mouse Desert hedgehog  
 CC gene. DNA encoding N-terminal sequences (see V62397) was obtained  
 CC by further PCR amplifications (see V62404-06). The invention  
 CC provides Desert hedgehog polynucleotides and protein, a monoclonal  
 CC antibody (Mab) that recognises the protein, a process for producing  
 CC the protein, and a method for detecting the protein using the Mab.  
 CC The hedgehog protein, DNA and Mab can be used to elucidate  
 CC hereditary morphological abnormalities in humans to establish their  
 CC treatments and diagnoses.  
 XX  
 SQ Sequence 1188 BP; 179 A; 376 C; 419 G; 214 T; 0 other;

alignment\_scores:  
 Quality: 949.00 Length: 179  
 Ratio: 5.302 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 99.441

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x V62395 ..

Align seg 1/1 to: V62395 from: 1 to: 1188

```

1  AAlaInserCySgLyProGlyArgLyProValGlyArgArgTyrAl 17
   |||||||
58  GCCCAGAGCTGGGGCGGGCGGGCGGCGTGGCGCGCGCGCTATGC 107
   |||||||
17  AArgLySgInLeuValProLeuLeuTyrLySgInPheValProGlyValP 34
   |||||||
108  GCGCAAGAGCTGCTGCCGCTACTACAGCAATTGTGGCGCGCTGC 157
   |||||||
34  roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50
   |||||||
158  CAGACCGGACCTGGGCGCCAGTAGGCGCGAGGGGGGTGGCGAAG 207
   |||||||
51  GlysSerGluArgPheArgAspLeuValProAsnTyrAsnProAspIleI 67
   |||||||
208  GGCCTCGAGCGCTTCGGGACCTCGTGCCCACTCAACCCGACATCAT 257
   |||||||
67  ePheLyAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgC 84
   |||||||
258  CTTCAAGAGTAGAGAACAGTAGAGCGCGCCCTGATGACCGCAACGTT 307
   |||||||
84  yLySgLyArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpPro 100
   |||||||
308  GTAAAGAACGGGTGAACGCTTTGGCCATTGCCGTATGAACTGTGGCCC 357
   |||||||
101  GLyValArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisHisAl 117
   |||||||
358  GGAGTGGCGCTACGAGTAGCTGAGGCTGGGACGAGGACGCCACACCC 407
   |||||||
117  aGlnAspSerLeuHisTyrGluGlyArgAlaLeuAspIleThrTrpSerA 134
   |||||||
408  TCAGGATTCACCTACACAGGAGCGCGCTTTGGACATCACTACCTCTG 457
   |||||||
134  sPArAspArgAsnLySTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150
   |||||||
458  ACCGGGACCGCAACAAGTAGTGGTGTGGCGCGCTCCGCAATGCAAGCC 507
   |||||||
151  GLyPheAspTrpValTyrTyrGluSerArgAsnHisValHisValSerVa 167
   |||||||
508  GGCTTCGACTGGCTACTACAGATCCCGCAACACATCCACCTGCGGT 557
   |||||||
167  lLyAlaAspAsnSerLeuAlaValArgAlaGly 179

```

seq\_name: /SIBS2/gcdata/geneseq/geneseq/NA1998.DAT:V62393

558 CAAGCTGATTAACACTGCGGTCGGGCGGCGGC 594

seq\_documentation\_block:  
 ID V62393 standard; cDNA; 528 BP.

AC V62393;  
 DT 02-FEB-1999 (first entry)

DE Human Desert hedgehog mature protein cDNA.

KW Desert hedgehog; HuDhH; human; ds.

OS Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT mat\_peptide 1..528  
 FT /\*tag= a

PN EP874048-A2.

PD 28-OCT-1998.

PE 24-APR-1998; 98EP-0303187.

PR 14-APR-1998; 98JP-0117873.

PR 25-APR-1997; 97JP-0121578.

XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

XX Ariyasu T, Nakamura S, Orita K;

XX WPI: 1998-544642/47.

DR P-PSDB: W79593.

PT Human Desert hedgehog protein - and corresponding DNA and monoclonal antibody

PS Claim 8; Page 20-21; 39pp; English.

XX This cDNA sequence codes for novel human Desert hedgehog protein  
 CC mature polypeptide (see W79593). Human Desert hedgehog protein  
 CC DNA (HuHh) was originally identified following extensive  
 CC screenings of human cell lines by PCR using primers (see V62402-03)  
 CC based on the mouse Desert hedgehog gene. These screenings  
 CC indicated that the ARH-77 (ATCC CRL-1621) cell line, which is  
 CC derived from the plasma cell of a leukemia patient, expressed a  
 CC specific gene at an elevated level. Sequencing confirmed it to  
 CC be a novel human gene showing homology to the mouse Desert hedgehog  
 CC gene. Nucleotide sequences (see V62394-95) encoding precursor  
 CC forms (see W79594-95) of human Desert hedgehog are also claimed,  
 CC as are a monoclonal antibody (Mab) that recognises the protein, a  
 CC process for producing the protein, and a method for detecting the  
 CC protein using the Mab. The hedgehog protein, DNA and Mab can be  
 CC used to elucidate hereditary morphological abnormalities in humans  
 CC to establish their treatments and diagnoses.

XX Sequence 528 BP; 101 A; 160 C; 181 G; 86 T; 0 other;

alignment\_scores:  
 Quality: 936.00 Length: 176  
 Ratio: 5.318 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 99.432

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x V62393 ..

Align seg 1/1 to: V62393 from: 1 to: 528

```

4 CysGlyProGlyArgGlyProValGlyValArgArgTyrAlaArgLysGI 20
1 TCGCGGCCGGCCGGGGCCGGTGGCGGGCGCTAGCGCCGCAAGCA 50
20 nLeuValProLeuLeuTyrLysGlnPheValProGlyValProGlyLysArgT 37
51 GCTGTGGCGCTACTCTACAGCAATTTGTGTCCGGCGGTGCAAGAGCGGA 100
37 hrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaAlaArgLysSerGlu 53
101 CCCGTGGCGCCAGTGGCGCCAGCGAGGAGCGGTGGCAAGCGCGTCCGAG 150
54 ArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLysAs 70
151 CGCTCCGGGACCTGTGCTCCACTACACCCGACATCATCTTCAAGGA 200
70 pGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysGluA 87
201 TGAGGAGAACAGTGGAGCGCCGCTGATGACCGAAGCTTGTAAGGAGAC 250
87 rGValAsnAlaLeuAlaIleAlaValMetAsnMetTTPProGlyValArg 103
251 GGGTGAACGGCTTTGGCCATTGCGGTGAAGAACATGTGGCCCGAGATGCC 300
104 LeuArgValThrGluGlyTTPAspGluAspGlyHisHisAlaGlnAspSe 120
301 CTAGAGAGTACAGGCTGGGAGAGGAGCCGACCCAGCTCAGGATTC 350
120 rLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgAspA 137
351 ACTCCACTACGAGAGCGCGTGGTGGACATCATACGCTGACCGGAGCC 400
137 rGAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPheAsp 153
401 GCAACAAGTAGAGGTGTCTGGCGGCTCGCAGTGAAGCGGCTTCGAC 450
154 TrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAlaAs 170
451 TGGGTCTACATGAGAGTCCCGAACACATCATGCTGCGTCAAACTGA 500
170 pAsnSerLeuAlaValArgAlaGlyGly 179
501 TAACTCACTGGCGGCTCCGGCGGCGGCG 528
seq_name: /studs2/gcgdata/geneseq/geneseqn/MA1998.DAT.V62394
seq_documentation_block:
ID V62394 standard; cDNA: 1122 BP.
AC V62394:
XX 02-FEB-1999 (first entry)
DT
XX
DE Human Desert hedgehog protein cDNA.
XX
KM Desert hedgehog; HuDHH; human; ds.
XX
OS Homo sapiens.
XX
FH key Location/Qualifiers
FT mat_peptide 1..528
FT /*tag= a
XX
XX EP874048-A2.
XX
XX 28-OCT-1998.
XX
XX 24-APR-1998: 98EP-0303187.
XX
XX 14-APR-1998: 98JP-0117873.
XX
XX 25-APR-1997: 97JP-0121578.
XX
XX (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

```

```

XX Ariyasu T, Nakamura S, Orita K;
XX WPI; 1998-544642/47.
XX P-PDB; W79594.
XX
XX Human Desert hedgehog protein - and corresponding DNA and monoclonal
XX antibody
XX
XX Claim 9; Page 22-23; 39pp; English.
XX
XX This cDNA sequence codes for a precursor (see W79594) of a novel
XX human Desert hedgehog protein (see also W79593). Human Desert
XX hedgehog protein DNA (HuDHH) was originally identified following
XX screenings of human cell lines by PCR using primers (see V62402-03)
XX based on the mouse Desert hedgehog gene. These screenings
XX indicated that the ARH-77 (ATCC CR1-1621) cell line, which is
XX derived from the plasma cell of a leukemia patient, expressed a
XX specific gene at an elevated level. Sequencing confirmed it to
XX be a novel human gene showing homology to the mouse Desert hedgehog
XX gene. DNA encoding C-terminal sequences (see V62398) was obtained
XX by further PCR amplifications (see V62407-08). The invention
XX provides Desert hedgehog polynucleotides and protein, a monoclonal
XX antibody (Mab) that recognises the protein, a process for producing
XX the protein, and a method for detecting the protein using the Mab.
XX The hedgehog protein, DNA and Mab can be used to elucidate
XX hereditary morphological abnormalities in humans to establish their
XX treatments and diagnoses.
XX
XX Sequence 1122 BP; 170 A; 351 C; 403 G; 198 T; 0 other;
XX
XX
XX alignment_scores:
XX      Quality: 936.00      Length: 176
XX      Ratio: 5.318      Gaps: 0
XX Percent Similarity: 100.000      Percent Identity: 99.432
XX
XX alignment_block:
XX US-08-900-220C-17_COPY_20_198 x V62394
XX
XX Align seg 1/1 to: V62394 from: 1 to: 1122
XX
XX 4 CysGlyProGlyArgGlyProValGlyValArgArgTyrAlaArgLysGI 20
XX 1 TCGCGGCCGGCCGGGGCCGGTGGCGGGCGCTAGCGCCGCAAGCA 50
XX 20 nLeuValProLeuLeuTyrLysGlnPheValProGlyValProGlyLysArgT 37
XX 51 GCTGTGGCGCTACTCTACAGCAATTTGTGTCCGGCGGTGCAAGAGCGGA 100
XX 37 hrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaAlaArgLysSerGlu 53
XX 101 CCCGTGGCGCCAGTGGCGCCAGCGAGGAGCGGTGGCAAGCGCGTCCGAG 150
XX 54 ArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLysAs 70
XX 151 CGCTCCGGGACCTGTGCTCCACTACACCCGACATCATCTTCAAGGA 200
XX 70 pGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysGluA 87
XX 201 TGAGGAGAACAGTGGAGCGCCGCTGATGACCGAAGCTTGTAAGGAGAC 250
XX 251 GGGTGAACGGCTTTGGCCATTGCGGTGAAGAACATGTGGCCCGAGATGCC 300
XX 87 rGValAsnAlaLeuAlaIleAlaValMetAsnMetTTPProGlyValArg 103
XX 301 CTAGAGAGTACAGGCTGGGAGAGGAGCCGACCCAGCTCAGGATTC 350
XX 120 rLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArgAspA 137
XX 351 ACTCCACTACGAGAGCGCGTGGTGGACATCATACGCTGACCGGAGCC 400

```

137 rgasnlstygylleuleualaargleualaValaGluAlaGlypheasp 153  
 401 GCAACAAAGTATGGGTGCTGGCGCCCTCGAGTGAAGCCGGCTTCGAC 450  
 154 TTPVALTYRTYRGLUSERARGASNHISVALHISVALSERVALYALAAS 170  
 451 TGGGTCTACTACGAGTCCCGCAACCAATCCACGTGCGTCAAAAGCTCA 500  
 170 PAsnSerLeuAlaValaArgAlaGlyly 179  
 501 TAACCTACTGGCGGTCCGGCGCGCGCGC 528  
 seq\_name: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:Q91642  
 seq\_documentation\_block:  
 ID Q91642 standard; cDNA; 1190 BP.  
 XX  
 AC Q91642;  
 XX  
 DT 18-MAR-1996 (first entry)  
 XX  
 DE Mouse desert hedgehog protein gene.  
 XX  
 KM Mouse; desert hedgehog gene; probe; primer; diagnostic;  
 KM nervous system disorder; gene therapy; antibody; ds.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT exon 1..303  
 FT /\*tag- a  
 FT exon 304..567  
 FT /\*tag- b  
 FT exon 568..1188  
 FT /\*tag- c  
 XX  
 PN WO9518856-A1.  
 XX  
 PD 13-JUL-1995.  
 XX  
 PE 30-DEC-1994; 94MO-US14992.  
 XX  
 PR 14-DEC-1994; 94US-0356060.  
 PR 30-DEC-1993; 93US-0176427.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY.  
 XX  
 PI Ingham PW, McMahon AP, Tabin CJ;  
 XX  
 DR WPI: 1995-255060/33.  
 DR P-PSDB: R77345.  
 XX  
 PT Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful  
 PT to treat degenerative nervous system disorder(s) and in gene  
 PT therapy.  
 XX  
 PS Claim 5; Page 135-37; 210pp; English.  
 CC The sequence encodes a mouse desert hedgehog protein, homologous  
 CC to a Drosophila hedgehog protein (R77337), and has been isolated by  
 CC low stringency screening of a mouse genome DNA library, in phage  
 CC lambda. The sequence contains 3 homologous regions, encoding a  
 CC single open reading frame interrupted by introns. Splicing has  
 CC been confirmed by polymerase chain reaction amplification of first  
 CC strand cDNA generated from adult testicle RNA. Probes and primers  
 CC derived from hedgehog sequences may be used as diagnostic agents  
 CC for neuromuscular, autonomic or central nervous system disorders,  
 CC and the gene may also be used in gene therapy. Antibodies  
 CC generated from the encoded protein may be used as therapeutic or  
 CC research reagents.

SQ Sequence 1190 BP; 194 A; 371 C; 399 G; 226 T; 0 other;  
 alignment\_scores:  
 Quality: 932.00 Length: 179  
 Ratio: 5.207 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 97.207  
 alignment\_block:  
 US-08-900-220c-17\_copy\_20\_198 x Q91642 ..  
 Align seg 1/1 to: Q91642 from: 1 to: 1190  
 1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgTyral 17  
 58 GCCAGAGCTGGCGGGCGCGGAGAGACCGGTTGGCGGGCGGATATGT 107  
 17 AArgGlyGlnLeuValProLeuLeuTyrtysGlnPheValProGlyValP 34  
 108 GCGCAAGCAACTTGTGCTGTATACAAGCAGTTTGTGCCAGTATGC 157  
 34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50  
 158 CCGAGCGGACCTCGGCGCGGAGTGGCCAGCGAGGAGGAGGTTAACAG 207  
 51 GlySerGluArgPheArgAspLeuValProAsnTyraPnProaspIleI 67  
 208 GGGTGGAGGCTTCGGGACCTGCTGATACCACTACACACCCGACATAT 257  
 67 ePheLysAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84  
 258 CTTCAGAGATGAGAGAGACAGCGCGCAGACCCCTGATGACAGAGCGTT 307  
 84 YstLysGluArgValAsnAlaLeuAlaIleAlaValMetasnmEtTrpPro 100  
 308 GCAAGACCGGGTGAACCTCTAGCATCGCGGTGATGAATGTGGCC 357  
 101 GlyValArgLeuArgValThrGluGlyTTPAspGluAspIleHisAlaI 117  
 358 GGAGTAGCGCTACGTGTGACTGAAAGCTGGGAGAGAGGCGCACACAGC 407  
 117 eGlnAspSerLeuHisTyrtGluArgAlaLeuAspIleThrThSera 134  
 408 ACAGGATTCACCTCCACTAGAAAGCCGTGCTGGACATCACACACCTCTG 457  
 134 sPArgAspArgAsnLysTyrtGlyLeuLeuAlaArgLeuAlaGluAla 150  
 458 ACCGTGACCGTATATAGTATGTTGTGGCGCGCTAGCTGTGAGAGCC 507  
 151 GlyPheAspTrpValTyrtTyrtGluSerArgAsnHisValHisValSerya 167  
 508 GGATTCGACGGGTCTACTAGAGTCCCGCAACCAATCCACATCCAGTATCG 557  
 167 llysalAspAsnSerLeuAlaValaArgAlaGlyly 179  
 558 CAAAGCTGATTAACCTACGTGCGGTCCGAGCGGAGGC 594

**THIS PAGE BLANK (USPTO)**

---

OM of: US-08-900-220c-17\_COPY\_20\_198 to: GenEmbl:\* out\_format : pfs  
Date: Apr 22, 2001 12:04 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

```
-MODEL=frame+pzr.model -DEV=xip  
-O=/cgn2_1/USPTO_spool/US08900220/runat_19042001_104047_21668/app-query.fasta_1.253  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -GAPOP=4.500  
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -GAPOP=6.000  
-GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=blomsu62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0  
-MAXLEN=2000000000 -USPR=US08900220_@CGNL_1_4706 -NCPU=6  
-ICPU=3 -LONLOG -NO_XLPPX -WAIT -THREADS=1
```

## Search information block:

Query: US-08-900-220c-17\_COPY\_20\_198

Query length: 179

Database: GenEmbl:\*

Database sequences: 1283235

Database length: -1216004940

Search time (sec): 1235.880000

## score\_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_pat1:AX054868	950.00	1409.25	2.8e-70	1190	AX054868 Sequence 15 from Pat
gb_pat1:AR021199	932.00	1382.48	8.7e-69	1190	AR021199 Sequence 3 from Paten
gb_pat1:AR063081	932.00	1382.48	8.7e-69	1190	AR063081 Sequence 2 from Paten
gb_pat1:AX054856	932.00	1382.48	8.7e-69	1190	AX054856 Sequence 3 from Paten
gb_r01:MM0HH	932.00	1382.48	8.7e-69	1191	X76292 M.musculus (129/Su) Dht
gb_ov:XL026349	823.50	1221.12	8.4e-60	1191	U26349 Xenopus laevis cephalic
gb_r01:MM085610	809.00	1199.51	1.3e-58	1197	U26350 Xenopus laevis hedgehog
gb_pat1:AR021198	743.50	1097.99	6.1e-53	1203	U85610 Mus musculus Indian hed
gb_pat1:AR063080	741.00	1097.91	6.2e-53	1277	AR063080 Sequence 1 from Paten
gb_pat1:AX054854	741.00	1097.91	6.2e-53	1277	AX054854 Sequence 1 from Paten
gb_ov:CHKLPAMD	741.00	1096.42	7.2e-53	1567	L28099 Gallus gallus Sequence
gb_pat1:AX054866	740.50	1095.42	8.5e-53	1622	AX054866 Sequence 13 from Pat
gb_pat1:AR021201	740.00	1096.22	7.6e-53	1313	AR021201 Sequence 7 from Paten
gb_pat1:AX054863	740.00	1096.22	7.6e-53	1313	AX054863 Sequence 4 from Paten
gb_pat1:AR063085	738.00	1092.65	1.2e-52	1425	AR063085 Sequence 6 from Paten
gb_r01:MM0SHH	738.00	1091.91	1.3e-52	1576	L38518 Homo sapiens sonic hedg
gb_pat1:AX054864	737.00	1091.75	1.4e-52	1314	AX054864 Sequence 11 from Pat
gb_ov:GGU58511	734.50	1085.12	3.2e-52	1958	U58511 Gallus gallus Indian he
gb_ov:NEMSPA	723.00	1081.99	4.7e-52	1807	D63339 Cynops pyrrhogaster son
gb_pat1:AX054858	723.00	1080.79	5.5e-52	1281	AX054858 Sequence 5 from Paten
gb_pat1:EL1679	723.00	1086.48	2.7e-52	530	EL1679 DNA encoding hedgehog p
gb_pat1:AR021202	723.00	1071.26	1.9e-51	1256	AR021202 Sequence 9 from Paten
gb_pat1:AR063084	723.00	1071.26	1.9e-51	1256	AR063084 Sequence 5 from Paten
gb_pat1:AX054862	723.00	1071.26	1.9e-51	1256	AX054862 Sequence 9 from Paten
gb_ov:DR030711	723.00	1069.12	2.5e-51	1684	U30711 Danio rerio sonic hedge
gb_ov:BRSONHP	723.00	1068.32	3.7e-51	2600	L235669 B.reticulo shh mRNA for sc
gb_ov:ZEFVHN1A	721.00	1065.95	2.7e-51	1251	L23585 Brachydanio rerio (vhn-1
gb_ov:DR030710	721.00	1068.32	2.7e-51	1251	U30710 Danio rerio tigray-winkl
gb_pat1:AX054870	718.00	1061.55	6.5e-51	1715	L23440 Rat (vhn-1) mRNA, comp
gb_r01:RATVHN1A	718.00	1061.55	6.5e-51	1715	U26314 Xenopus laevis sonic he
gb_ov:XL026314	716.00	1059.34	8.7e-51	1546	L35248 Xenopus laevis sonic he
gb_ov:AB029748	716.00	1058.37	9.6e-51	1718	AB029748 Paralichthys olivaceu
gb_ov:ELKSHS	715.00	1055.98	1.3e-50	2447	L35213 Xenopus laevis morphog
gb_ov:AF047466	713.00	1057.44	1.1e-50	1635	AF047466 Notophthalmus virides
gb_ov:XL026404	713.00	1056.54	1.2e-50	1230	U26404 Xenopus laevis banded h
gb_ov:AF003532	707.00	1053.37	1.9e-50	1549	AF003532 Pleurodeles waltl son
gb_int1:AB044709	707.00	1041.20	8.9e-50	2963	AB044709 Gyrillus blinaculatus m

gb\_ov:DRECHINDA + 704.00 1041.81 8.2e-50 1479 Y08426 D.reticulo mRNA for EHR  
gb\_r01:AF148226 + 660.00 986.18 1.0e-46 386 AF148226 Rattus norvegicus d  
gb\_int2:BFAMPHIHH + 640.00 940.12 3.8e-44 3610 Y13858 Branchiostoma florida  
gb\_ov:AF144100 + 618.00 923.15 3.3e-43 417 AF144100 Bos taurus sonic he

seq\_name: gb\_pat1:AX054868

## seq\_documentation\_block:

LOCUS AX054868 1190 bp DNA PAT 13-JAN-2001  
DEFINITION Sequence 15 from Patent WO0073337.  
ACCESSION AX054868  
VERSION AX054868.1 GI:12228297  
KEYWORDS  
SOURCE human.  
ORGANISM human.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

1 (bases 1 to 1190)

## AUTHORS

Peplinsky, R.B., Taylor, F. and Garber, E.

## TITLE

Polymer conjugates of hedgehog proteins and uses

## JOURNAL

Patent: WO 0073337-A 15 07-DEC-2000;  
BIOGEN, INC. (US)

## FEATURES

Location/Qualifiers

## source

1..1190  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 176 a 375 c 424 g 215 t  
ORIGIN

## alignment\_scores:

Quality: 950.00 Length: 179  
Ratio: 5.307 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x AX054868 ..

Align seg 1/1 to: AX054868 from: 1 to: 1190

1 AAGInserCgCgGlyProGlyArgGlyProValAGlyATGATGATGAT 17  
|||||  
58 GCCCAGAGCTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 107  
17 arglyGgInleuValProleuLeuTyrTgGInPheValProGlyVal 34  
|||||  
108 GCGCAGCGAGCTGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 157  
34 roGlyATGThrLeuAGlyAlaSerGlyProAlaGlyAGlyValAlaArg 50  
|||||  
158 CAGAGCGAGCGCTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 207  
51 GlySerGlyATGThrLeuArgAspLeuValProAsnTyrAsnProAsp 67  
|||||  
208 GGCCTCGAGCGCTCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 257  
67 ePheTyrAspGlyGluGluAsnSerGlyAlaAspArgLeuMetThrGly 84  
|||||  
258 CTTCAAGAGTGGAGAGACAGTGGAGCCAGCCAGCCAGTATGACGAG 307  
308 GCAAGAGAGAGGAGTGAAGCTTGGCCATGCGCGTGTGAACATGGG 357  
101 GYVALATGLeuATGValThrGluGlyTTPArgGluAspGlyNH2HisAl 117  
|||||  
358 GAGAGTGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 407  
117 aGlnAspSerLeuNH2TyrGluGlyAlaGlyAlaLeuAspLeuThrTyr 134  
|||||  
408 TCAGAGTATGACGACCACTAGAGAGCGGCGGCTTTGGACATCAGTCTG 457  
134 sPARGAspArgAsnTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150

|||||  
458 ACCGCGACCCCAACAGTATGGGTGCTGCGCGCTCGCAGTGAAGCC 507  
151 G1YPheSPTRPValTYrTYGluSerARgsnHsValHsValSerVa 167  
|||||  
508 GGCTTCAGCTGGGTCTACTAGAGTCCCGCAGCAGCTCCAGCTGGGT 557  
167 lYsAlAspAsnSerLeuAlaValArGAlaGlyGly 179  
|||||  
558 CAAAGCTGATTAACCTACTGCGCGCTCGCGCGCGCG 594

seq\_name: gb\_pat1:AR021199

seq\_documentation\_block:  
LOCUS AR021199 1190 bp DNA PAT 05-DEC-1998  
DEFINITION Sequence 3 from patent US 5789543.  
ACCESSION AR021199  
VERSION AR021199.1 GI:3975814  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1190)  
AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.  
TITLE Vertebrate embryonic pattern-inducing proteins and uses related thereto  
JOURNAL Patent: US 5789543-A 3 04-AUG-1998;  
FEATURES  
source 1. 1190 Location/Qualifiers  
BASE COUNT 194 a 371 c 399 g 226 t  
ORIGIN

alignment\_scores:  
Quality: 932.00 Length: 179  
Ratio: 5.207 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 97.207

alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x AR021199 ..

Align seg 1/1 to: AR021199 from: 1 to: 1190

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgArgTYRAl 17  
|||||  
58 GCCAGAGAGCTGCGCGCGCGCGCGAGCAGCGTTGCGCGCGCTTATGT 107  
17 aArgYGLnLeuValProLeuLeuTYrLYSGlnPheValProGlyValP 34  
|||||  
108 GCGCAACCACTTGCTGCTGCTATACAAAGCAGTTGTGCGCGAGTATGC 157  
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50  
|||||  
158 CCGAGCGGACCTTGCGCGCGAGTGGCCAGCGAGGAGGATACAAAG 207  
51 GlySerGluArgPheArgAspLeuValProAsnTYrAsnProAspIleI 67  
|||||  
208 GGCTCGAGAGCGCTTCCGGACCTGCTACCACTAACCAACCCGACATAAT 257  
67 ePheLYsAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84  
|||||  
258 CTTCAAGGATGAGGAGAAAGCGCGCGAGACCGCTGATGACAGAGGCTT 307  
84 YsLYSGluArgValAlaSerGlyAlaLeuAlaIleAlaValMetAsnMetTrpPro 100  
|||||  
308 GCAAGAGCGGGGTGAACGCTCTAGCCATCGGGGTGATGAACATGTGCCC 357  
101 G1YValArgLeuArgValAlaThrGluGlyTRPAspGluAspGlyHsHsAl 117  
|||||  
358 GGANtAGCGCTACTGTGACTGAAGGCTGGAGCAGCAGCCCAACCGCC 407  
117 aGlnArgSerLeuHsTYrGluGlyArgAlaLeuAspIleThrThrSerA 134

|||||  
408 ACAGATTAACCTCCACTACAGAGGCCGCTGCTTGACATACCAACGCTG 457  
134 sPArgAspArgAsnLYsTYrGlyLeuLeuAlaArgLeuAlaValGAla 150  
|||||  
458 ACCGTGACCGTAAATTAATATAGTTTGTGGCGCGCCTAGCTGTGAAGCC 507  
151 G1YPheSPTRPValTYrTYGluSerARgsnHsValHsValSerVa 167  
|||||  
508 GGATTCAGCTGGGTCTACTAGAGTCCCGCAGCAGCTCCAGCTGGGT 557  
167 lYsAlAspAsnSerLeuAlaValArGAlaGlyGly 179  
|||||  
558 CAAAGCTGATTAACCTACTGCGCGCTCGCGCGCGAGGC 594

seq\_name: gb\_pat1:AR063081

seq\_documentation\_block:  
LOCUS AR063081 1190 bp DNA PAT 29-SEP-1999  
DEFINITION Sequence 2 from patent US 5844079.  
ACCESSION AR063081  
VERSION AR063081.1 GI:5990772  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1190)  
AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.  
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related thereto  
JOURNAL Patent: US 5844079-A 2 01-DEC-1998;  
FEATURES  
source 1. 1190 Location/Qualifiers  
BASE COUNT 194 a 371 c 399 g 226 t  
ORIGIN

alignment\_scores:  
Quality: 932.00 Length: 179  
Ratio: 5.207 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 97.207

alignment\_block:

US-08-900-220C-17\_COPY\_20\_198 x AR063081 ..

Align seg 1/1 to: AR063081 from: 1 to: 1190

1 AlaGlnSerCysGlyProGlyArgGlyProValGlyArgArgArgTYRAl 17  
|||||  
58 GCCAGAGAGCTGCGCGCGCGCGCGAGCAGCGTTGCGCGCGCTTATGT 107  
17 aArgYGLnLeuValProLeuLeuTYrLYSGlnPheValProGlyValP 34  
|||||  
108 GCGCAACCACTTGCTGCTGCTATACAAAGCAGTTGTGCGCGAGTATGC 157  
34 roGluArgThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArg 50  
|||||  
158 CCGAGCGGACCTTGCGCGCGAGTGGCCAGCGAGGAGGATACAAAG 207  
51 GlySerGluArgPheArgAspLeuValProAsnTYrAsnProAspIleI 67  
|||||  
208 GGCTCGAGAGCGCTTCCGGACCTGCTACCACTAACCAACCCGACATAAT 257  
67 ePheLYsAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgC 84  
|||||  
258 CTTCAAGGATGAGGAGAAAGCGCGCGAGACCGCTGATGACAGAGGCTT 307  
84 YsLYSGluArgValAlaSerGlyAlaLeuAlaIleAlaValMetAsnMetTrpPro 100  
|||||  
308 GCAAGAGCGGGGTGAACGCTCTAGCCATCGGGGTGATGAACATGTGCCC 357  
101 G1YValArgLeuArgValAlaThrGluGlyTRPAspGluAspGlyHsHsAl 117

358 GGAGTACGCTACGTGTGACTGAAAGCTTGAGACGAGGACGGCCACACAGC 407  
|||||  
117 aglnAspSerLeuHisTyrGluGlyValArgAlaLeuAspIleThrPheSerA 134  
|||||  
408 ACAGAGATTCACTCCACTACGAGAGCGCGTCTTGGACATCACACAGCTTG 457  
|||||  
134 spArgAspArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAla 150  
|||||  
458 ACCGTGACCGCTAAATAAGTATGTTTGTGGCGCGCTACGCTGTGGAGCC 507  
|||||  
151 GlyPheAspTyrValTyrTyrGluSerArgAsnHisValHisValSerVal 167  
|||||  
508 GGATTTCGACTGGGCTCTACTACGAGTCCCGCAACCAACATCCACAGTTCGT 557  
|||||  
167 llyslAlaAspAsnSerLeuAlaValArgAlaGlyGly 179  
|||||  
558 CAAAGCTGATTACTACTGAGGGGTCCGAGACCGGAGAGC 594  
|||||

seq\_name: gb\_pat1:AX054856





104 LeuArGValThrGlUGlyTTPAspGluAspGlyNHSHSLAGlnAspSe 120  
 364 CTACGGGTTACGAGGGTTGGGACGATGACACACGACGTCACGACAC 413  
 120 rLeuNHSTyrgLUGlyArGAlaLeuAspLleThrHrSerAspArGAspA 137  
 414 TCTGCATTTATGAAGGCGCGCACTGCACATTCACATCTGACCGAGACA 463  
 137 rGAsnTystrGlyLeuLeuAlaArGLeuAlaValGlnAlaGlyPheAsp 153  
 464 GGAACAATAACGCGATGTGGCAGCTGCACATGGAAGCTGGCTTGGAC 513  
 154 TTPValTyrTyrgLUGlySerArGAsnHISValHISValSerValLysAlaAs 170  
 514 TGGGTATATATGAATCCAAAGCACATCATGTCTGTGTAAGACAGA 563  
 170 pAsnSerLeuAlaValArGAlaGlyGly 179  
 564 TAACCTCCCTGGGCTGCTGCTGCTGCTG 591

seq\_name: gb\_ov:XLU26350

seq\_documentation\_block: 1197 bp mRNA VRT 05-JAN-1996  
 LOCUS XLU26350  
 DEFINITION Xenopus laevis hedgehog protein 4 (hh4) mRNA, complete cds.  
 ACCESSION U26350  
 VERSION U26350.1 GI:1147829  
 KEYWORDS  
 SOURCE African clawed frog.  
 ORGANISM  
 Xenopus laevis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 Xenopodinae; Xenopus.  
 1 (bases 1 to 1197)  
 Eker, S.C., McGrew, L.L., Lai, C.J., Lee, J.J., von Kessler, D.P.,  
 Moon, R.T. and Beachy, P.A.  
 Distinct expression and shared activities of members of the  
 hedgehog gene family of Xenopus laevis  
 Development 121 (8), 2337-2347 (1995)  
 2 (bases 1 to 1197)  
 Eker, S.C., McGrew, L.L., Lai, C.L., Lee, J.J., von Kessler, D.P.,  
 Moon, R.T. and Beachy, P.A.  
 Direct Submission  
 Submitted (04-MAY-1995) Stephen C. Eker, Molecular Biology and  
 Genetics, Johns Hopkins University, 725 N. Wolfe St./714 PCIB,  
 Baltimore, MD 21205, USA

FEATURES  
 source  
 1. .1197  
 /organism="Xenopus laevis"  
 /db\_xref="taxon:8355"  
 gene  
 1. .1197  
 /gene="hh4"  
 /gene="hh4"  
 /note="related to Xenopus laevis cephalic hedgehog,  
 encoded by GenBank Accession Number U26349; secreted  
 signaling molecule; Method: conceptual translation  
 supplied by author"  
 /codon\_start=1  
 /product="hedgehog protein 4"  
 /protein\_id="AA85164.1"  
 /db\_xref="GI:1147830"  
 /translation="MPAVRLLIAACCCMLLPVRCGPGRGPGVGGRRRYRRLVPL  
 LTKQFVNPVPEKTKGASGSGKTRGSERTKLVPNPNPILFDEKNTADRLMTE  
 RCKRDVALAIVNMMWPKLRLVTEGWDGSHAHSLHTEGRALDITSDRKNKY  
 GMLARLAEAGFDWVYMSKAHIVSVNTDSSLVRSQSGCPGTAMVMMEGKKPLS  
 ELKLGDTFTDEGLIHISVLLFHRDPYKATFVLEAGHPTRKLLVPHNHLFI  
 KSSSTGQPFPEAYRVOIGDLIYVNTGVOSSKVVAVSDEQGYVAPMPEGTL  
 VDGYLTSCYATVESHTLAHSLARLFLQGLASMLPDLHTSDGVNMYCHILYLAKEY  
 LMDMP"  
 CDS  
 295 a 286 c 314 g 302 t

BASE COUNT

ORIGIN

Alignment\_scores:  
 Quality: 809.00 Length: 177  
 Ratio: 4.787 Gaps: 2  
 Percent Similarity: 95.480 Percent Identity: 84.746

Alignment\_block:  
 US-08-900-220c-17\_copy\_20\_198 x XLU26350 ..  
 Align seg 1/1 to: XLU26350 from: 1 to: 1197

4 CysGlyProGlyArGlyProVal...GlyArGArGArGlyAlaArgly 19  
 70 TGTGTCCTAGCGCGGGGGCGGTGGCGGCGCGCGGTACTAGTAGAG 119  
 19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGluA 36  
 120 A...CTGGTCCCACTGCTTTACAAAGCATGTTGTGCTTAATGTGCCAGAGA 166  
 36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArGValAlaArGlySer 52  
 167 AGACTCTAGGGCGCACTGGCACTCAGAGGCAAGATCCGACGGGCTTG 216  
 53 GluArGpHeArGAspLeuValProAsnTyrAsnProAspLleLlePheLy 69  
 217 GAAAGTTCATCAAGTTGGTCCCACTCAACCCAGACATCATCTTCAA 266  
 69 sAspGluGluAsnSerGlyAlaAspArGLeuMetThrGluArGysLysG 86  
 267 GGATGAGAGAACACACAGAGCTGACCTCATCATGATGAGCGCTGCAGAG 316  
 86 LuArGValAsnAlaLeuAlaIleAlaValMetArGMetThrProGlyVal 102  
 317 ACAGACTCATAGCTTGGCTATCTGTGATGACATGAGCCAGGCTTG 366  
 103 ArGLeuArGValThrGlUGlyTTPAspGluAspGlyNHSHSLAGlnAs 119  
 367 AACCTCGGGGTGACGAGGGCTGGATGAGAGGCGCACACGCTCAGCA 416  
 119 pSerLeuNHSTyrgLUGlyArGAlaLeuAspLleThrHrSerAspArG 136  
 417 CTCTTTGCATTTATGAAGGCGCGCACTGCACATTCACATCTGACCGAG 466  
 136 sPrArGsnTystrGlyLeuLeuAlaArGLeuAlaValGlnAlaGlyPhe 152  
 467 ACAGAACAAATATGTCATGTCGCCGCTGACATGGAAGCTGGCTT 516  
 153 AspTTPValTyrTyrgLUGlySerArGAsnHISValHISValSerValLysAl 169  
 517 GACTGGGTGATTTACGATCCAAAGCACATCATGTCTGTGTAAGAC 566  
 169 aAspAsnSerLeuAlaValArGAlaGlyGly 179  
 567 AGATAACTCTCCCTGGTGTGCTGCTGCTGCTG 597

seq\_name: gb\_r01:MM085610

seq\_documentation\_block:  
 LOCUS MM085610 2103 bp mRNA ROD 16-MAR-1997  
 DEFINITION Mus musculus Indian hedgehog protein (Ihh) mRNA, complete cds.  
 ACCESSION U85610  
 VERSION U85610.1 GI:1890096  
 KEYWORDS  
 SOURCE house mouse.  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 2103)  
 Valentini, R.P., Brookhiser, W.T., Parr, J.J., Yang, T., Briggs, J.,  
 Dressler, G. and Holzman, L.B.  
 Post-translational Processing and Renal Expression of Mouse Indian

Hedgehog  
J. Biol. Chem. (1997) In press  
2 (bases 1 to 2103)  
REFERENCE  
AUTHORS  
Valentini, R.P., Brookhiser, W.T., Park, J., Yang, T., Briggs, J.,  
Dressler, G. and Holzman, L.B.  
TITLE  
JOURNAL  
Direct Submission  
Submitted (15-JAN-1997) Internal Medicine/Division of Nephrology,  
University of Michigan Medical School, 1560 MSRB II--Box 0676, Ann  
Arbor, MI 48109-0676, USA  
FEATURES  
Location/Qualifiers  
1. 2103  
/organism="Mus musculus"  
/strain="CD-1"  
/db\_xref="taxon:10090"  
184..1533  
/gene="Inh"  
184..1533  
/gene="Inh"  
/codon\_start=1  
/product="Indian hedgehog protein"  
/protein\_id="AA049692.1"  
/db\_xref="GI:1890097"  
/translation="MESPRAQTQPSPKLSOPRAHLAHOAPSALPGYPAMPAL  
RRLRRCLELLILLIPARAGCGRGRVRSRRPRKLYPLAYKQSPVPRKTCAS  
GREESKIASSESRFKELTPNYPDITFDEENTGADRLTQCKDRLNSLAISNMOM  
PEKLVNTEGWDSDHSESLYHGEKRAVDITSDRNKYGILRLAVEAGFDWVY  
ESKAHVHGSVKSEHSAAKTGGCFPAGAVLENERVALSAVKGEDRLAMGEDGP  
TFSVDLIEIDREPNRLRAFOVETODPPRLALTPAHLLFIIDNHTPEPAHRAFTAS  
HYQPOYVIVSGVPGIOPARVAIVSTHVALGSAPLTRHGLTVEDVVAISCPAAVADH  
HIAOLAFLMPLRFPSILAMGSMTPSEGVHMYPMIYRLGLLLESTFHLIGMSGAS"

BASE COUNT 376 a 684 c 628 g 415 t  
ORIGIN

alignment\_scores:  
Quality: 743.50 Length: 179  
Ratio: 4.479 Gaps: 1  
Percent Similarity: 92.737 Percent Identity: 75.419

alignment\_block:  
US-08-900-220C-17\_COPY\_20\_198 x MM085610 ..  
Align seg 1/1 to: MM085610 from: 1 to: 2103

```
1  Alaglnsercysgilyproglyarglyprovalglyargargtyral 17
||||:||||:||||:||||:||||:||||:||||:||||:||||:
370 GCGCGGGGCTGGCGGGCGGGCGG...GTGGTGGCAGCGCGCGAGAGCC 416
17 aarglyglnleuvalproleuutyrlsglnphevalproglyvalp 34
417 GCCTGCGAAGCTGTGCTCTTGCCTACAGACATTCAGCCCAAGCTGC 466
34 rogluatgrhleuglyalasercglyproalagluclyargvalalaarg 50
467 CGGAGAGACCCCTGGCGCGCGCGGCGCTACGAGCAAGATGCGCGCC 516
51 glysercyluargpheargaspleuvalproasnTyraNproaspilleil 67
517 AGCTCTAGGCGCTTCAAGAGCTCACCCCAACTACATATCCGACATCAT 566
67 ephelyaspsgluglnasercglyalaaspargleumethrcgluargc 84
567 CTTCAAGAGACGAGAGACAGGGGTGCGACCGCTCATGACCCAGCGCT 616
84 yslsgluargvalasnalaleualailealaValaMetasmettrpPro 100
617 GCAAGAGACGCTGTCACTGACCTCTGTGTCATGATGAACACGCGCT 666
101 gllyvalargleuargvalThrgluglyTTPaspcluaspglyshisAl 117
667 GGTGTGAAGTGGCGGTGACCGAGAGCTGGGATGAAGATGCGCATCACTC 716
117 aglinspserleuthistyrgluclyargalaleuaspllethrthrsra 134
```

```
717 AGAGAGCTCTTACATATGAGGGCGCGCGTGGATATACACCACTCAG 766
134 sspaGAPARGAsnlystYrclyleuAlaArgleuAlaValaGlAla 150
767 ACCGTGACCGAAATAGATATGACTGCTGGCGCGCTTACGATGGAGCC 816
151 GlypheasprtpvalYtyrTgluSerArgAsnHisValHisValSera 167
817 GGCTTCAGCTGGGTGTATACGAGTCCAGGCCACGCTGCTCTGT 866
167 llyAlaAspasSerleuAlaValaArgAlaGlyly 179
867 CAAGCTGACCATTCGGCGCGCTGCCAAGACAGGTGC 903
```

seq\_name: gb\_pat1:AR021198

seq\_documentation\_block: 1277 bp DNA PAT 05-DEC-1998  
LOCUS AR021198  
DEFINITION Sequence 1 from patent US 5789543.  
ACCESSION AR021198  
VERSION AR021198.1 GI:3975813  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1277)  
AUTHORS Ingham, P.W., McMahon, A.P. and Tablin, C.J.  
TITLE Vertebrate embryonic pattern-inducing proteins and uses related  
thereto US 5789543-A 1 04-AUG-1998:  
JOURNAL Location/Qualifiers  
ORIGIN 1. 1277  
BASE COUNT 262 a 396 c 387 g 232 t  
ORIGIN

alignment\_scores:  
Quality: 741.00 Length: 177  
Ratio: 4.411 Gaps: 2  
Percent Similarity: 94.915 Percent Identity: 74.011

alignment\_block:  
US-08-900-220C-17\_COPY\_20\_198 x AR021198 ..  
Align seg 1/1 to: AR021198 from: 1 to: 1277

```
3 sercysgilyproglyarglyprovalglyargargtyralaargly 19
||||:||||:||||:||||:||||:||||:||||:||||:||||:
76 ACTGTGGACCGAGGAGGGC...ATTGGAAAAAGAGGCGACCCCAAAA 122
19 sglleuvalproleuutyrlsglnphevalproglyvalproglua 36
123 G...CTGACCCCTTACCTATAGCAGTTATTCCTCAATGTGCACAGNA 169
36 rgrthleuglyalasercglyproalagluclyargvalalaarglyser 52
170 AGACCTTAGGGGCGCAGTGAAGATATGAAGAGATCACAAGAACTCC 219
53 gluatgrpheargaspleuvalproasnTyraNproaspillelephely 69
220 GAGGAGATTTAAGAACTACACCAATTTACAACTTATATTTTAA 269
69 saspgluginasercglyalaaspargleumethrcgluargcylasg 86
270 GATGAAGAGACACCGGAGCTGCACAGACTGATGACTCAGCGCTCAAGG 319
86 luargvalasnalaleualailealaValaMetasmettrpProglyVal 102
320 ACAAGCTGAATGCGCTGCGATCTCGGTGATGAACAGTGGCGGGGTG 369
103 ArgleuargvalThrgluglyTTPaspcluaspglyshisAlaGlnas 119
```

```
370 AAGCTGGGGGACCGAGGCGCTGGAGCAGAGATGGCCATCATCTCCGAGGA 419
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
119 pserleuhtstygltgluylaralaleuaspilletthrtserasparga 136
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
420 ATCGTGACACTACGAGGGTCCGCCGTGCATCACCACGTGGATCCGG 469
136 spargasnlystygltyleuleualaargleualaialgluaglyphe 152
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
470 ACCGAGAGAGTACGAGATGCTGGCCGCTCCGTCGAGGCGGCTTC 519
153 asptpvaltyttrgluserargasnhtsvalhsvalservallysal 169
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
520 GACTGGGTCTACTACGAGTCCAGGCGCACATCCACTGCTCCGTAAGC 569
169 aaspasnserleualaialargalaglygly 179
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
570 AGAAACTCAGTGGCGCAGCAATCAGAGAGC 600
```

seq\_name: gb\_pat1:AR063080

seq\_documentation\_block:

LOCUS AR063080 1277 bp DNA PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5844079.

ACCESSION AR063080

VERSION AR063080.1 GI:5990771

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1277)

AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.

TITLE Vertebrate embryonic pattern-inducing proteins, and uses related

thereto

JOURNAL Patent: US 5844079-A 1 01-DEC-1998;

FEATURES

source 1..1277

BASE COUNT 262 a 396 c 387 g 232 t

ORIGIN

alignment\_scores:  
Quality: 741.00 Length: 177  
Ratio: 4.411 Gaps: 2  
Percent Similarity: 94.915 Percent Identity: 74.011

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x AR063080 ..

Align seg 1/1 to: AR063080 from: 1 to: 1277

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyAlaArgly 19
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
76 ACTTGTGGACGACGAGGGGC...ATTGGAAAAAGAGGACCCCAAAAA 122
19 scInleuValProLeuLeuTyLysGlnPheValProGlyValProGlu 36
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
123 G...CTGACCCCGTTAGCCTATAGCAGATTATTCCTCATGTGGCAGAA 169
36 rgtHrleuGlyAlaSerGlyProAlaGlyGlyArgValAlaArgGlySer 52
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
170 AGACCTAGGGGCGCAGTGGAGATATGAAGGAGATCACAAGAAATCC 219
53 GluArgPheArgAspLeuValProAsnTyraProAspIleIlePhe 69
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
220 GAGAGATTAAAGAACTAACCCCAATTATCAACCCGACATTATTTTAA 269
69 sasptgluGluAsnSerGlyAlaAspArgLeuMetHrGlyArgCysLysG 86
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
270 GGATGAGAGAAACAGCGGAGCTGACGAGACTGATGACTACGCGCTGAAG 319
86 luArgValAsnAlaLeuAlaIleAlaValaMetAsnMetTrpProGlyAl 102
```

```
320 ACAACCTGAATGCCCTGGGATCTCGGTATGAACAGTGGCGCGGGTG 369
103 ArgLeuAlaGlyValThrGlyLysArgGlyAspGlyHisAlaGlnAs 119
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
370 AAGCTGGGGGTACCGAGGGCTGGGACGACGAGATGCCCATCATCTCCGAGGA 419
119 pserleuhtstygltgluylaralaleuaspilletthrtserasparga 136
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
420 ATCGTGACACTACGAGGGTCCGCCGTGCATCACCACGTGGATCCGG 469
136 spargasnlystygltyleuleualaargleualaialgluaglyphe 152
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
470 ACCGAGAGAGTACGAGATGCTGGCCGCTCCGTCGAGGCGGCTTC 519
153 asptpvaltyttrgluserargasnhtsvalhsvalservallysal 169
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
520 GACTGGGTCTACTACGAGTCCAGGCGCACATCCACTGCTCCGTAAGC 569
169 aaspasnserleualaialargalaglygly 179
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
570 AGAAACTCAGTGGCGCAGCAATCAGAGAGC 600
```

seq\_name: gb\_pat1:AX054854

seq\_documentation\_block:

LOCUS AX054854 1277 bp DNA PAT 13-JAN-2001

DEFINITION Sequence 1 from Patent WO0073337.

ACCESSION AX054854

VERSION AX054854.1 GI:12228290

KEYWORDS

SOURCE

ORGANISM

REFERENCE 1 (bases 1 to 1277)

AUTHORS Pepinsky, R.B., Taylor, F. and Garber, E.

TITLE Polymer conjugates of hedgehog proteins and uses

JOURNAL Patent: WO 0073337-A 1 07-DEC-2000;

FEATURES

source 1..1277

BASE COUNT 262 a 396 c 387 g 232 t

ORIGIN

alignment\_scores:  
Quality: 741.00 Length: 177  
Ratio: 4.411 Gaps: 2  
Percent Similarity: 94.915 Percent Identity: 74.011

alignment\_block:

US-08-900-220c-17\_COPY\_20\_198 x AX054854 ..

Align seg 1/1 to: AX054854 from: 1 to: 1277

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyAlaArgly 19
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
76 ACTTGTGGACGACGAGGGGC...ATTGGAAAAAGAGGACCCCAAAAA 122
19 scInleuValProLeuLeuTyLysGlnPheValProGlyValProGlu 36
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
123 G...CTGACCCCGTTAGCCTATAGCAGATTATTCCTCATGTGGCAGAA 169
36 rgtHrleuGlyAlaSerGlyProAlaGlyGlyArgValAlaArgGlySer 52
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
170 AGACCTAGGGGCGCAGTGGAGATATGAAGGAGATCACAAGAAATCC 219
53 GluArgPheArgAspLeuValProAsnTyraProAspIleIlePhe 69
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
220 GAGAGATTAAAGAACTAACCCCAATTATCAACCCGACATTATTTTAA 269
```

Percent Similarity: 94.915 Percent Identity: 74.011  
 alignment\_block:  
 US-08-900-220C-17\_COPY\_20\_198 x CHKZPAMED ..  
 Align seg 1/1 to: CHKZPAMED from: 1 to: 1567

```

3 SerCysGlyProGluArgGlyProValGlyArgArgArgTyrAlaArgLy 19
302 ACCTGTGGACGACGAGGAGG...ATTGAAAAAGGAGGACCCCAAAA 348
19 sGlnLeuValProLeuLeuTyrLysGlnPheValProGlyValProGlu 36
349 G...CTGACCCCGTTAGCCCTATAGCGATTATTTCCAAATGTGGCAGAGA 395
36 rGhrLeuGlyAlaSerGlyProAlaGlyArgValAlaArgGlySer 52
396 AGACCCCTAGGGGCCAGTGGAAATATGAAGGAAGATCACAACTCC 445
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePhe 69
446 GAGGATTTTAAAGAACTAACCCCAATTACACCCCTGACATTTATTTAA 495
69 sAspGluGluAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
496 GGATGAAGAGAACACGGGAGCTGACAGCTGATGACTACACGCTGCAGAG 545
86 LuArgValAlaAlaLeuAlaIleAlaValMetAsnMetProGlyVal 102
546 ACAAGCTGAATGCCCTGGCATCTCGGTGATGAACCAAGTGGCCGGGGTG 595
103 ArgLeuArgValThrGluGlyTyrPaspGluAspGlyHisAlaGlnAs 119
596 AAGCTGGGGTGTGACCGAGGCTGGGACAGATGGCCATCACTCCGAGGA 645
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrSerAspArg 136
646 ATCCCTCACTACGAGGGTCCGCCCTGGGACATCCACACGTCGGATGGG 695
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
696 ACCGACCAAGTACGGAATGCTGGCCGCCCTCGCGTGCAGGGCGGCTTC 745
153 AspTyrValIlyTyrGluSerArgAsnHisValHisValSerValIysAl 169
746 GACTGGCTCTACTACGATCCCAAGCGCCACATCCATCGTCCGCAAGC 795
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
796 AGAAACTCACTGCGACGCAATCGAGAGGC 826
seq_name: gb_pat1:AX054866
  
```

```

seq_documentation_block: 1567 bp mRNA VRT 04-OCT-1999
LOCUS CHKZPAMED
DEFINITION Gallus gallus sonic hedgehog mRNA, complete cds.
ACCESSION L28099
VERSION L28099.1 GI:453526
KEYWORDS
SOURCE
ORGANISM Gallus gallus
            chicken.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
            1 (bases 1 to 1567)
REFERENCE Riddle, R.D., Johnson, R.L., Laufer, E. and Tabin, C.
            Sonic hedgehog mediates the polarizing activity of the ZPA
            JOURNAL Cell 75, 1401-1416 (1993)
MEDLINE 94094333
FEATURES
            location/Qualifiers
            1..1567
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            /rissue_type="embryo"
            /dev_stage="stage 20-22"
            227..1504
            /codon_start=1
            /product="sonic hedgehog"
            /protein_id="AA072428.1"
            /db_xref="GI:453527"
            /translation="MVEMLLRLLVGFICALLVSSGLTGGPGRGIGKRHRPKLTP
            LAYKOFIYNVAEKTLGASRGYEGKLTIRNSERKELTPVNPDIIRKDEENQADRLMT
            ORCKDKLAIALISVNMQPGVKLRTEMDMDEHSGESLSHHEGABDITSSDPRRSK
            YGMRLARLAVENGEPDVIYTESKAHICSVKAESVAKSGCCPGSATYHLEHGTRKLY
            KDLSPGDVNLADADGRLLYSDLEFLTDRMSSRLFLYIETRPRLRLITTAHLFL
            VAPHNQSTGTSQALFASNVKPGRGVVLGGGQQLPASVHVSLLREASGAY
            APLTAQGTILINRVASCYAVIEESHMAHMAAPRLAOGILALCPDCAIPTAATT
            TGIHWYSRLYRIGSWVLDGALHPLGMVAPAS"
  
```

```

CDS
            1..1567
            /organism="Gallus gallus"
            /db_xref="taxon:9031"
            /rissue_type="embryo"
            /dev_stage="stage 20-22"
            227..1504
            /codon_start=1
            /product="sonic hedgehog"
            /protein_id="AA072428.1"
            /db_xref="GI:453527"
            /translation="MVEMLLRLLVGFICALLVSSGLTGGPGRGIGKRHRPKLTP
            LAYKOFIYNVAEKTLGASRGYEGKLTIRNSERKELTPVNPDIIRKDEENQADRLMT
            ORCKDKLAIALISVNMQPGVKLRTEMDMDEHSGESLSHHEGABDITSSDPRRSK
            YGMRLARLAVENGEPDVIYTESKAHICSVKAESVAKSGCCPGSATYHLEHGTRKLY
            KDLSPGDVNLADADGRLLYSDLEFLTDRMSSRLFLYIETRPRLRLITTAHLFL
            VAPHNQSTGTSQALFASNVKPGRGVVLGGGQQLPASVHVSLLREASGAY
            APLTAQGTILINRVASCYAVIEESHMAHMAAPRLAOGILALCPDCAIPTAATT
            TGIHWYSRLYRIGSWVLDGALHPLGMVAPAS"
BASE COUNT 347 a 457 c 475 g 288 t
ORIGIN
  
```

alignment\_scores:  
 Quality: 741.00 Length: 177  
 Ratio: 4.411 Gaps: 2

```

seq_documentation_block: 1622 bp DNA PAT 13-JAN-2001
LOCUS AX054866
DEFINITION Sequence 13 from Patent WO0073337.
ACCESSION AX054866
VERSION AX054866.1 GI:12228296
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            1 (bases 1 to 1622)
REFERENCE Pepinsky, R.B., Taylor, F. and Garber, E.
            Polymers conjugates of hedgehog proteins and uses
            JOURNAL Patent: WO 0073337-A 13 07-DEC-2000;
            BIOGEN, INC. (US)
FEATURES
            location/Qualifiers
            1..1622
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
BASE COUNT 277 a 549 c 510 g 286 t
  
```



AUTHORS Ingham, P.W., McMahon, A.P. and Tabin, C.J.  
TITLE Vertebrate embryonic pattern-inducing proteins, and uses related  
thereto

JOURNAL Patent: US 5844079-A 4 01-DEC-1998;

FEATURES Location/Qualifiers  
source 1..1313

BASE COUNT 254 a 411 c 428 g 220 t

ORIGIN

alignment\_scores:

Quality: 740.00 length: 177  
Ratio: 4.431 Gaps: 2  
Percent Similarity: 94.350 Percent Identity: 74.576

alignment\_block:

US-08-900-220c-17\_copy\_20\_198 x AR063083 ..

Align seg 1/1 to: AR063083 from: 1 to: 1313

```
3 SerCysGlyProGlyArgGlyProValGlyArgArgArgTyrAlaArgLy 19
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
70 GCGTGTGGGCGCGGCGAGGGG...TTTGAAAGAGCGCGCACCCCAAAA 116
19 SGInLeuValProLeuLeuTyrTrpGlnPheValProGlyValProGluA 36
| ||:::||||| |||||:::|||||:::|||||:::|||||:::|||||
117 G...CTGACCCCTTACCTACAGCAGTTATTTCCACAGTACCGCAGA 163
36 rGThrLeuGlyAlaSerGlyProAlaGluGlyArgValAlaArgGlySer 52
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
164 AGACCCAGCGCGCGCGCGAGATATGAGGAGGAGATCACAGAACTCC 213
53 GluArgPheArgAspLeuValProAsnTyrAsnProAspIleIlePheLy 69
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
214 GAACGATTTAAGGAACACACCCCAATTACACCCGACATCATATTAA 263
69 sAspGluGlnAsnSerGlyAlaAspArgLeuMetThrGluArgCysLysG 86
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
264 GGATGAGGAAACACGCGGAGCAGACCGGCTGATGACTCAGAGGTGCAAG 313
86 LuArgValAsnAlaLeuAlaIleAlaValMetAsnMetTrpProGlyVal 102
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
314 ACAAGTTAAATGCCCTTGCCATCTCTGTGATGAACCACTGCCCTGAGTG 363
103 ArgLeuArgValThrGluGlyTyrTrpAspGluAspGlyHisAlaGlnAs 119
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
364 AGGCTGCGAGTGAACCGAGCGCTGGAGTGAAGCGGCATTCAGAGGA 413
119 pSerLeuHisTyrGluGlyArgAlaLeuAspIleThrThrSerAspArg 136
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
414 GTCTTAACATAAGGCTGAGCAGTGAATCACCACGTCGCCGCGGG 463
136 sPArgAsnLysTyrGlyLeuLeuAlaArgLeuAlaValGluAlaGlyPhe 152
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
464 ACCGCAACAGTACGCGATGCTGGCTGCCCTGCTGGAACGAGTTTC 513
153 AspTrpValTyrTyrGluSerArgAsnHisValHisValSerValLysAl 169
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
514 GACTGGGTACTATGAAATCCAAAGCTCACATCCACTGTCTGTGAAGC 563
169 aAspAsnSerLeuAlaValArgAlaGlyGly 179
|||||:::|||||:::|||||:::|||||:::|||||:::|||||
564 AGAAGACTCGGTGGCGCAAAATCCGGCGG 594
```